

# Northumbria Research Link

Citation: Dimmock, A. P., Welling, D. T., Rosenqvist, L., Forsyth, C., Freeman, M. P., Rae, Jonathan, Viljanen, A., Vandegriff, E., Boynton, R. J., Balikhin, M. A. and Yordanova, E. (2021) Modeling the Geomagnetic Response to the September 2017 Space Weather Event Over Fennoscandia Using the Space Weather Modeling Framework: Studying the Impacts of Spatial Resolution. *Space Weather*, 19 (5). e2020SW002683. ISSN 1542-7390

Published by: American Geophysical Union

URL: <https://doi.org/10.1029/2020sw002683> <<https://doi.org/10.1029/2020sw002683>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/46845/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

# Space Weather



## RESEARCH ARTICLE

10.1029/2020SW002683

### Key Points:

- Higher space weather modeling framework (SWMF) spatial resolution increased the variability of virtual ground magnetometer and inter-station differences
- By increasing the SWMF spatial resolution, the ability to model geomagnetically induced currents (GICs) in terms of amplitudes and timing can be improved
- Regardless of the spatial resolution, substorms remain a problem and are a crucial process to driving GICs

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

A. P. Dimmock,  
[andrew.dimmock@irfu.se](mailto:andrew.dimmock@irfu.se)

### Citation:

Dimmock, A. P., Welling, D. T., Rosenqvist, L., Forsyth, C., Freeman, M. P., Rae, I. J., et al. (2021). Modeling the geomagnetic response to the September 2017 space weather event over Fennoscandia using the space weather modeling framework: Studying the impacts of spatial resolution. *Space Weather*, 19, e2020SW002683. <https://doi.org/10.1029/2020SW002683>

Received 16 NOV 2020

Accepted 23 APR 2021

© 2021. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Modeling the Geomagnetic Response to the September 2017 Space Weather Event Over Fennoscandia Using the Space Weather Modeling Framework: Studying the Impacts of Spatial Resolution

A. P. Dimmock<sup>1</sup> , D. T. Welling<sup>2</sup> , L. Rosenqvist<sup>3</sup> , C. Forsyth<sup>4</sup> , M. P. Freeman<sup>5</sup> , I. J. Rae<sup>4,6</sup> , A. Viljanen<sup>7</sup> , E. Vande-griff<sup>2</sup>, R. J. Boynton<sup>8</sup> , M. A. Balikhin<sup>8</sup> , and E. Yordanova<sup>1</sup> 

<sup>1</sup>Swedish Institute of Space Physics, Uppsala, Sweden, <sup>2</sup>Department of Physics, University of Texas at Arlington, Arlington, TX, USA, <sup>3</sup>Swedish Defence Research Agency, Stockholm, Sweden, <sup>4</sup>UCL Mullard Space Science Laboratory, Dorking, UK, <sup>5</sup>British Antarctic Survey, Cambridge, UK, <sup>6</sup>Northumbria University, Newcastle upon Tyne, UK, <sup>7</sup>Finnish Meteorological Institute, Helsinki, Finland, <sup>8</sup>Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK

**Abstract** We must be able to predict and mitigate against geomagnetically induced current (GIC) effects to minimize socio-economic impacts. This study employs the space weather modeling framework (SWMF) to model the geomagnetic response over Fennoscandia to the September 7–8, 2017 event. Of key importance to this study is the effects of spatial resolution in terms of regional forecasts and improved GIC modeling results. Therefore, we ran the model at comparatively low, medium, and high spatial resolutions. The virtual magnetometers from each model run are compared with observations from the IMAGE magnetometer network across various latitudes and over regional-scales. The virtual magnetometer data from the SWMF are coupled with a local ground conductivity model which is used to calculate the geoelectric field and estimate GICs in a Finnish natural gas pipeline. This investigation has led to several important results in which higher resolution yielded: (1) more realistic amplitudes and timings of GICs, (2) higher amplitude geomagnetic disturbances across latitudes, and (3) increased regional variations in terms of differences between stations. Despite this, substorms remain a significant challenge to surface magnetic field prediction from global magnetohydrodynamic modeling. For example, in the presence of multiple large substorms, the associated large-amplitude depressions were not captured, which caused the largest model-data deviations. The results from this work are of key importance to both modelers and space weather operators. Particularly when the goal is to obtain improved regional forecasts of geomagnetic disturbances and/or more realistic estimates of the geoelectric field.

**Plain Language Summary** Geomagnetically induced currents (GICs) affect large ground-based conducting infrastructure and are associated with the dynamic behavior of geospace electrical currents that drive rapid variations of the surface geomagnetic field. Through geomagnetic induction with the ground conductivity, a geoelectric field is set up which causes unwanted currents to flow in large-scale technological systems. This can result in damage, which can eventually lead to failures and disruptions. To mitigate against GICs we need the capability to predict geomagnetic field variations at the surface with sufficient accuracy, knowledge of the ground conductivity, and a realistic description of the affected system. In this study, we use a global magnetohydrodynamic model (with additional integrated components) to model the surface geomagnetic field variations for the September 2017 event. We compare the simulated ground magnetic fields with those measured at equivalent locations. The spatial resolution of the model is modified to determine if this provides improved performance for capturing spatially structured geomagnetic disturbances. The modeled geomagnetic fields are employed with a ground conductivity model to compute GICs in a natural gas pipeline, which is compared with GIC recordings. We see that higher spatial resolution runs can improve GIC predictions when not driven by substorm activity.

## 1. Introduction

Space weather is increasingly recognized as a significant socio-economic risk to society due to our dependence on space- and ground-based technology, which can be adversely affected by the response of near-Earth space to extreme driving conditions. Of high importance is the geomagnetic induction problem where the dynamics of geospace currents cause large amplitude and rapid surface geomagnetic field perturbations. These can lead to geomagnetically induced currents (GICs) that flow in large conductive systems such as power lines and pipelines. As a result, GICs have been investigated for many years (Anderson et al., 1974; Kappenman, 2003; Lehtinen & Pirjola, 1985; Pulkkinen et al., 2015; Viljanen, 1997). To achieve a high degree of situational awareness and successfully mitigate GICs, we must build a high-level of physical knowledge of the magnetospheric system and translate that into numerical models. For actionable information, these numerical models should be able to reproduce the surface ground magnetic field perturbations with sufficient accuracy in terms of magnitude, polarity, and spatial structure. Numerical models also provide additional benefits of investigating past-events (e.g., Ngwira et al., 2013, 2014; Raeder et al., 2001) and tracing the source of surface perturbations to the magnetosphere. However, new studies have reaffirmed that geomagnetic perturbations are highly structured and complex (Ngwira et al., 2015, 2018; Pulkkinen et al., 2015), which poses significant physical and technical challenges from a modeling perspective; considering that physics-based models used for real-time geomagnetic predictions fall into the category of global magnetohydrodynamic (MHD). Undoubtedly, there remain unanswered questions on how well these models can capture spatially structured events and be applied to model GICs, which are driven by complex spatiotemporal geomagnetic disturbances.

The goal of the present study is to focus explicitly on the role of global MHD spatial resolution using the space weather modeling framework (SWMF) in capturing regional geomagnetic variations and modeling GICs at Fennoscandian latitudes between 60° and 70°. Although studies such as Pulkkinen et al. (2010) have shown that increasing the spatial resolution can improve performance, there are still unanswered questions. In this paper we aim to build on these past studies and shed light on the following: (1) does higher resolution always provide better performance at resolving geomagnetic disturbances? (2) how do extremely high-resolution runs of almost 8 million cells compare to lower resolution runs typically used in operational settings? (3) can we capture more regional variability using higher spatial resolution? (4) how are (1–3) affected by substorms, and (5) can higher resolution runs be beneficial to modeling highly complex events? To achieve this, we will present a case study of a simulation of the September 2017 event using the SWMF at an extremely high spatial resolution, which we compare with lower resolution runs and observations. In addition, we assess the regional variability using a similar method to Dimmock et al. (2020), while considering substorms in our analysis by systematically identifying events using ground magnetometer data and comparing them to the model results.

We selected this event since the period between September 7 and 8, 2017 was particularly active from a space weather context, resulting in global geomagnetic disturbances which have received deserved attention over the past few years (Clilverd et al., 2018; Dimmock et al., 2019; Piersanti et al., 2019). A key factor that contributed to the interest in this event was that the upstream conditions originated from the interaction between multiple Interplanetary Coronal Mass Ejections (ICMEs; Werner et al., 2019), which created a highly complex period of external driving in terms of shocks, turbulence, IMF rotations, and plasma fluctuations. This event is ideal for addressing (1–5) listed above since Fennoscandia measured significant extended geomagnetic disturbances (Dimmock et al., 2019), multiple substorms were reported, and it is expected to pose a challenge to numerical models.

Many relevant previous studies have compared magnetometer measurements with simulated ground magnetic fields (Kwagala et al., 2020; Pulkkinen et al., 2010, 2011, 2013; Shao et al., 2002; Yu & Ridley, 2008) using a variety of global MHD models. Shao et al. (2002) utilized the Lyon-Fedder-Mobarry (LFM) which showed reasonable agreement but tended to underestimate the observations during active periods. Also, there were small spikes (of 10–15 min duration) in the ground observations which were not captured, and the authors attributed this to the possible presence of auroral arcs that are not explicitly included in the model. Yu and Ridley (2008) performed a validation of the SWMF using ground magnetometers for seven events with varying degrees of strength in terms of auroral and equatorial indices. The authors primarily used the metrics of root mean squared difference and prediction efficiency to evaluate the model-data comparison. Overall

the model performance was good, but there were notable differences as a function of event, latitude, and season. However, the data contained more structure than the model, and the model tended to miss some localized large amplitude disturbances. Notably, the model magnetic fields were generally lower than the data, but this was not always systematic, meaning that in some cases, both the magnitude and ionospheric current patterns were distorted. Despite that, the AL (amplitude lower) index comparison showed that although the model underrepresented the data, the trend was similar.

Several papers (see Pulkkinen et al., 2010, 2011, 2013) have extensively explored the performance of global MHD models in capturing ground magnetic fields, which provide the framework for similar studies that build on these results. Pulkkinen et al. (2010) compared the ground magnetic field across multiple MHD models and also investigated the impacts of spatial resolution, which is also one of the goals in the present study. They observed an increase in performance for higher spatial resolution according to a minimum resolution of  $1/4 R_E$  and 2 million grid cells. The current study will significantly build on this result by using extremely high-resolution runs of 8 million cells and a minimum spatial resolution of  $1/16 R_E$ . We provide new perspectives on this topic by considering directly the effects during and outside of substorm periods on the ability to model GIC from simulation outputs. How this may affect the spatial structure of geomagnetic disturbances is also addressed.

Pulkkinen et al. (2011) compared 14 physics-based and empirical models across four events, where the performance was assessed based on multiple validation metrics. There was no systematic “best” model since the performance depended heavily on the metric and an event-by-event basis. Therefore it is clear that continual investigations of global MHD models for space weather purposes (which utilize new events) is highly beneficial to obtain a greater understanding of the applicability of these models to certain situations, which is a motivation for the present study. In general, most of the models tended to underestimate the maximum absolute horizontal magnetic field perturbation and corresponding  $dB_h/dt$ . A point was made that the best solution may be to choose from a selection of models that are optimized for various metrics. In other words, the appropriate model should be chosen based on the characteristics of the signal which are important to a given task or end-user. Pulkkinen et al. (2013) continued the community-wide validation of space weather models to assess the capabilities of the numerous models to forecast  $d\mathbf{B}/dt$  since this quantity is a reliable proxy for the geoelectric field (Viljanen et al., 2001). Although the models do not capture the instantaneous  $d\mathbf{B}/dt$ , the amplitudes at high-latitudes at the beginning of events were reproduced and therefore show some agreements. Thus, the models for such cases may be applicable given a specific forecast window. Still, it was concluded that localized predictions remain a significant challenge, which will also be considered in this study using even higher high-fidelity runs.

A consensus from the above-mentioned studies is that ring current models (see also Rae et al., 2010), spatial resolution, and ionospheric conductance (see also Mukhopadhyay et al., 2020) are all key elements in improving ground magnetic field forecasts from global MHD simulations. D. T. Welling et al. (2017) offered some new perspectives on the work by Pulkkinen et al. (2013), particularly that ionospheric conductance codes employed by MHD models may fall outside their range of validity for some inputs during strong-extreme events. The authors also pointed out that success in predicting  $\Delta B$  does not necessarily translate to  $d\mathbf{B}/dt$  which in general seems to be significantly under-predicted. In fact, Tóth et al. (2014) showed that the skill of forecasting  $d\mathbf{B}/dt$  (from SWMF) can be improved by indirectly adopting an empirical relationship between  $d\mathbf{B}/dt$  and the magnitude of  $\Delta\mathbf{B}$ , rather than computing it directly from  $d\mathbf{B}/dt = (\Delta\mathbf{B}(t+1) - \Delta\mathbf{B}(t)) / \Delta t$ . Thus, it is clear that addressing the ability to capture regional variations of  $\Delta\mathbf{B}$  during different events, locations, and under various magnetospheric conditions is highly beneficial to the feasibility of these models to the geoelectric field and GIC applications.

Earlier studies discussed (Anderson et al., 1974; Boteler & Jansen Van Beek, 1999) the spatial variability of geomagnetic disturbances, and recently, localized peaks in the geoelectric field and rate-of-change of the field ( $d\mathbf{B}/dt$ ) have been identified (Ngwira et al., 2015, 2018; Pulkkinen et al., 2015). Regional geomagnetic predictions remain challenging, and substorms (Akasofu, 1964) are one process that has been proposed to be capable of producing such regional variability (Ngwira et al., 2018). They are associated with dynamic ionospheric currents, correspondingly strong geomagnetic fields (Engebretson et al., 2019; Pulkkinen et al., 2003; Viljanen, Pulkkinen, et al., 2006), and may manifest as localized disturbances (Tsurutani

et al., 2015). For that reason, we consider the impacts of multiple substorms on the ability to model geomagnetic disturbances and how this affects GIC modeling capabilities.

Dimmock et al. (2020) recently performed a statistical study of the regional variability of  $d\mathbf{B}/dt$ , and we will adopt a similar methodology in this study to quantify the regional variations of  $\Delta B$ . In this study, the authors showed that  $d\mathbf{B}/dt$  significantly varies over hundreds of kilometers, and this has an impact on modeled GICs given a transmission line of comparable length. Diurnal trends in these data also suggested substorms may play a significant role (Freeman et al., 2019), further emphasizing the need to quantitatively assess the modeling capability of these events. Having said that, other events such as storm commencements are also capable of driving similar behavior (Smith et al., 2019). These investigations have spurred interest in assessing the ability of the current state-of-the-art models to predict  $d\mathbf{B}/dt$  and some regional variability of geomagnetic perturbations. The September 2017 event exhibited multiple substorms measured over Fennoscandia, which we will use to address the significance of capturing substorms on geomagnetic field predictions and determine how this may inhibit GIC modeling.

With the current interest in understanding spatially structured geomagnetic disturbances and the renewed interest in how state-of-the-art models may capture these effects, the key aspect of this study is that we perform three separate SWMF simulations with “low,” “medium,” and “very high” spatial resolutions, much higher than has been used in many previous studies. Our intention is not to perform model-validation as in previous studies but to explicitly investigate how the spatial resolution can impact our ability to model large-scale geomagnetic disturbances, smaller-scale depressions resulting from substorms, and the regional (few hundred km) geomagnetic variations measured between magnetometer stations. Finally, we quantitatively assess how this affects the accuracy of GIC modeling. Since no current studies have explicitly addressed these questions with such a high-resolution and using this complex event, we believe the results will be of interest to both the scientific community and also forecasters who may be considering increasing the spatial resolution of their models.

## 2. Data, Models, and Methods

### 2.1. IMAGE Data

We employ observations from the IMAGE magnetometer array (Tanskanen, 2009) to characterize the geomagnetic response over Fennoscandia. The network consists of 41 stations which are maintained by a collaboration of eight institutes located in Finland, Germany, Norway, Poland, Russia, and Sweden. Stations cover a geographical latitudinal range from 51° to 79° over the Baltic and Fennoscandian regions. This distribution proves ideal for studying the strength and variability of auroral currents. A map of the IMAGE stations is shown in Figure 1a.

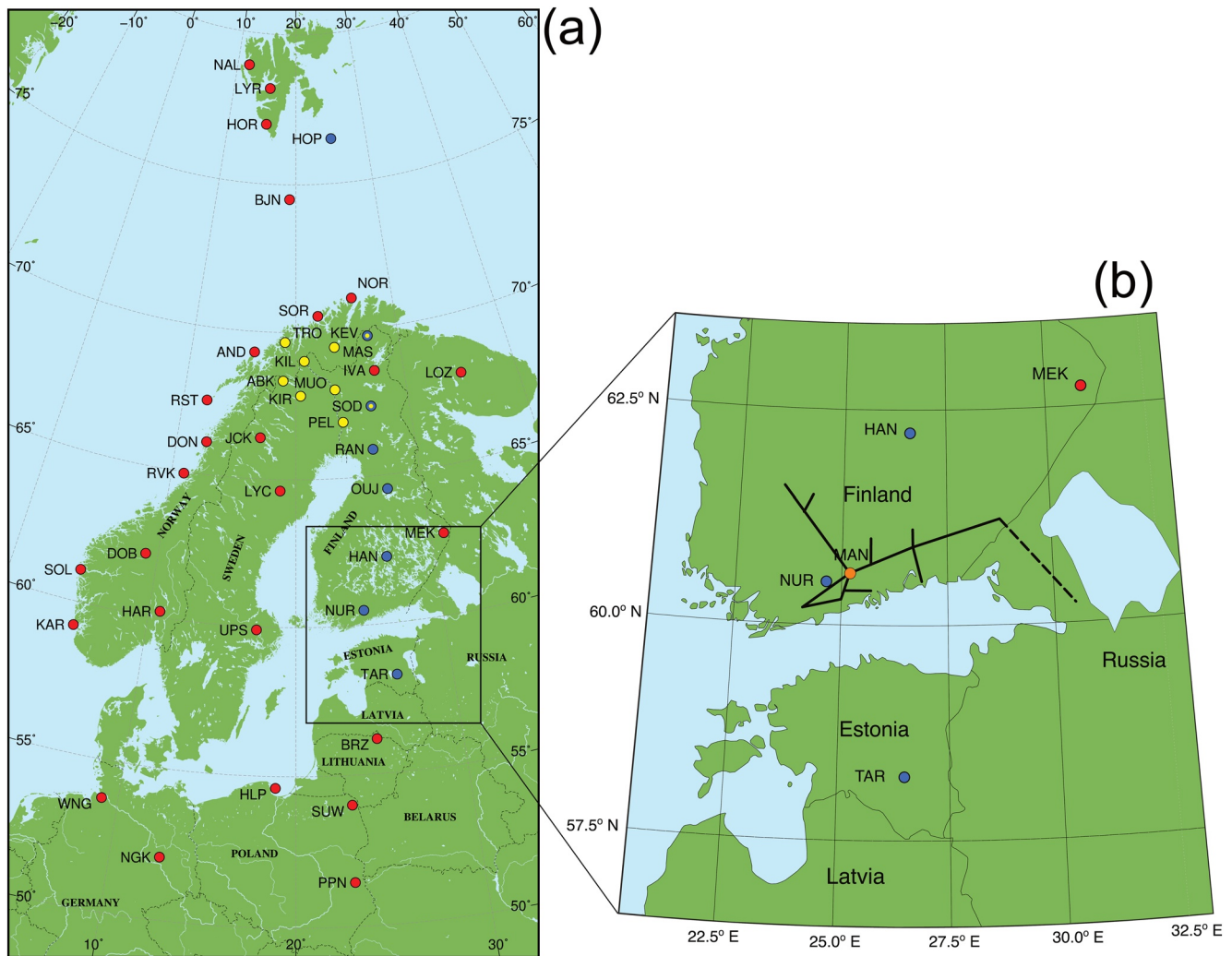
The temporal resolution of the IMAGE network is 10 s, meaning it can capture small temporal variations. Also, the dense inter-station separation at latitudes between 65° and 70° enables studies of spatial variations on the order of hundreds of kilometers. Significant spatio-temporal behavior of the auroral ionospheric currents is common during strong storms and thus, these two capabilities are crucial. Note that IMAGE data are in geodetic coordinates and are used throughout this paper. Therefore, the virtual magnetometer data output in geomagnetic coordinates is transformed to geodetic co-ordinates to be consistent with the measured magnetometer data.

In the present study, these stations are used in three manners: (1) to study the latitudinal behavior of  $\Delta B$ , (2) investigate spread of  $\Delta B$  over regional (hundreds of kilometers) scales, and (3) as inputs to a ground conductivity model. For the first, we exploit the notable latitudinal spread and employ the stations shown by the blue markers in Figure 1a. For the second, we extract a subset from the IMAGE network over a region where the stations are dense, shown by the yellow dots. Stations with both blue and yellow are used in both. Table 1 lists all the stations used and their geographic coordinates.

### 2.2. GIC Recordings in Mäntsälä

Since November 1998, GIC recordings have been conducted at the Finnish natural gas pipeline in Mäntsälä located at 60.6°, 25.2° geographic latitude and longitude, respectively. This places Mäntsälä approximately





**Figure 1.** Slightly modified from Dimmock et al. (2019). Panel (a) shows a map of the IMAGE magnetometer network. The blue dots correspond to the stations spread in latitude but with similar longitudes. The yellow dots show a subset of stations that are closely-spaced and used to analyze regional effects. Markers with blue and yellow are used for both purposes. All the stations used are listed in Table 1. Shown in panel (b) is the topology of the Finnish natural gas pipeline. The orange dot is placed at Mäntsälä, which is the location of the compressor station and where GICs are measured.

40 km from the Nurmijärvi magnetometer station. Figure 1b shows the spatial extent of this pipeline which is mainly comprised of eastward and northward main lines and some smaller branches. GICs are determined based on measuring the magnetic field directly above the pipeline, which is a combination of natural variations and those by currents flowing in the pipeline. The Nurmijärvi measurement is removed to eliminate the fields not originating from GICs, meaning the GICs can be computed according to Biot-Savart law. See Pulkkinen, Pirjola et al. (2001); Pulkkinen, Viljanen et al. (2001) for a more complete description of this procedure, and Dimmock et al. (2019) for a recent application. As pointed out by Pulkkinen, Pirjola et al. (2001); Pulkkinen, Viljanen et al. (2001), the signal of the cathodic corrosion protection adopted by the pipeline operators provides some measure of keeping the pipeline at a lower potential than the soil unavoidably gets mixed with the true GIC signal. Although this can cause some uncertainty, the measured GIC data that is a combination of these two signals has been proven to act as a true GIC (e.g., Viljanen, Pulkkinen, et al. 2006; Viljanen, Tanskanen, et al., 2006) since it is intrinsically related to  $dB/dt$  and the modeled geoelectric field. As shown by numerous studies, when geomagnetic variations are small, then so is the GIC ( $<1$  A) regardless of the presence of noise from the protection system.

Although measured locally, GICs are the combined effect from non-local contributions at many points within a system. Therefore, ideally, a line integral of the geoelectric field is performed over the system topology

**Table 1**  
The Geographic Latitude and Longitude of the IMAGE Stations That Were Used in This Study

Station	Code	Geogr. lat (°)	Geogr. lon (°)
Hopen Island <sup>b,y</sup>	HOP	76.51	25.01
Kevo <sup>b,y</sup>	KEV	69.76	27.01
Tromsø <sup>y</sup>	TRO	69.66	18.94
Masi <sup>y</sup>	MAS	69.46	23.70
Kilpisjärvi <sup>y</sup>	KIL	69.06	20.77
Abisko <sup>y</sup>	ABK	68.35	18.82
Muonio <sup>y</sup>	MUO	68.02	23.53
Kiruna <sup>y</sup>	KIR	67.84	20.42
Sodankylä <sup>b,y</sup>	SOD	67.37	26.63
Pello <sup>y</sup>	PEL	66.90	24.08
Ranua <sup>b</sup>	RAN	65.90	26.41
Oulujärvi <sup>b</sup>	OIJ	64.52	27.23
Hankasalmi <sup>b</sup>	HAN	62.25	26.60
Nurmijärvi <sup>b</sup>	NUR	60.50	24.65
Tartu <sup>b</sup>	TAR	58.26	26.46

*Note.* These stations are also highlighted by the blue and yellow markers in the IMAGE map shown in Figure 1 and noted by the b and y superscripts.

(Pulkkinen, Pirjola, et al., 2001; Pulkkinen, Viljanen, et al., 2001) since it is the geoelectric field parallel to a conductive line which is important (Ingham & Rodger, 2018). However, in this case, only local magnetic field measurements are available, and the magnetic field at other locations is unknown. Still, the Mäntsälä GIC can be modeled assuming that the geoelectric field is uniform over the spatial scale of the pipeline. Although this can be violated during the presence of highly localized ionospheric currents, for this event it was demonstrated that currents were of comparable scale to the pipeline (Dimmock et al., 2019). Using the ground conductivity values provided by Viljanen, Pulkkinen, et al. (2006); Viljanen, Tanskanen, et al. (2006) (resistivity of 0.02567 S/m to a depth of 150 km and then 2.5974 S/m to infinite depth), we can derive the geoelectric field directly from the measured/modeled local magnetic field. Based on the favorable agreement between the geoelectric field produced by 1D and 3D conductivity models, this geoelectric field should be a reasonable approximation of the region as shown by Dimmock et al. (2019). This comparison indicates that this location did not contain significant conductivity gradients. According to Pulkkinen, Pirjola et al. (2001); Pulkkinen, Viljanen et al. (2001), the horizontal geoelectric field ( $E_x$ ,  $E_y$ ) can be used to calculate the GIC based on the following expression

$$GIC(t) = aE_x(t) + bE_y(t), \quad (1)$$

At Mäntsälä,  $a$  and  $b$  are  $-70$  Akm/V and  $88$  Akm/V, respectively. Although these parameters were not determined recently, and these parameters will not exactly match the current situation. Having said that, it has been shown recently by Dimmock et al. (2019) and also in this study that

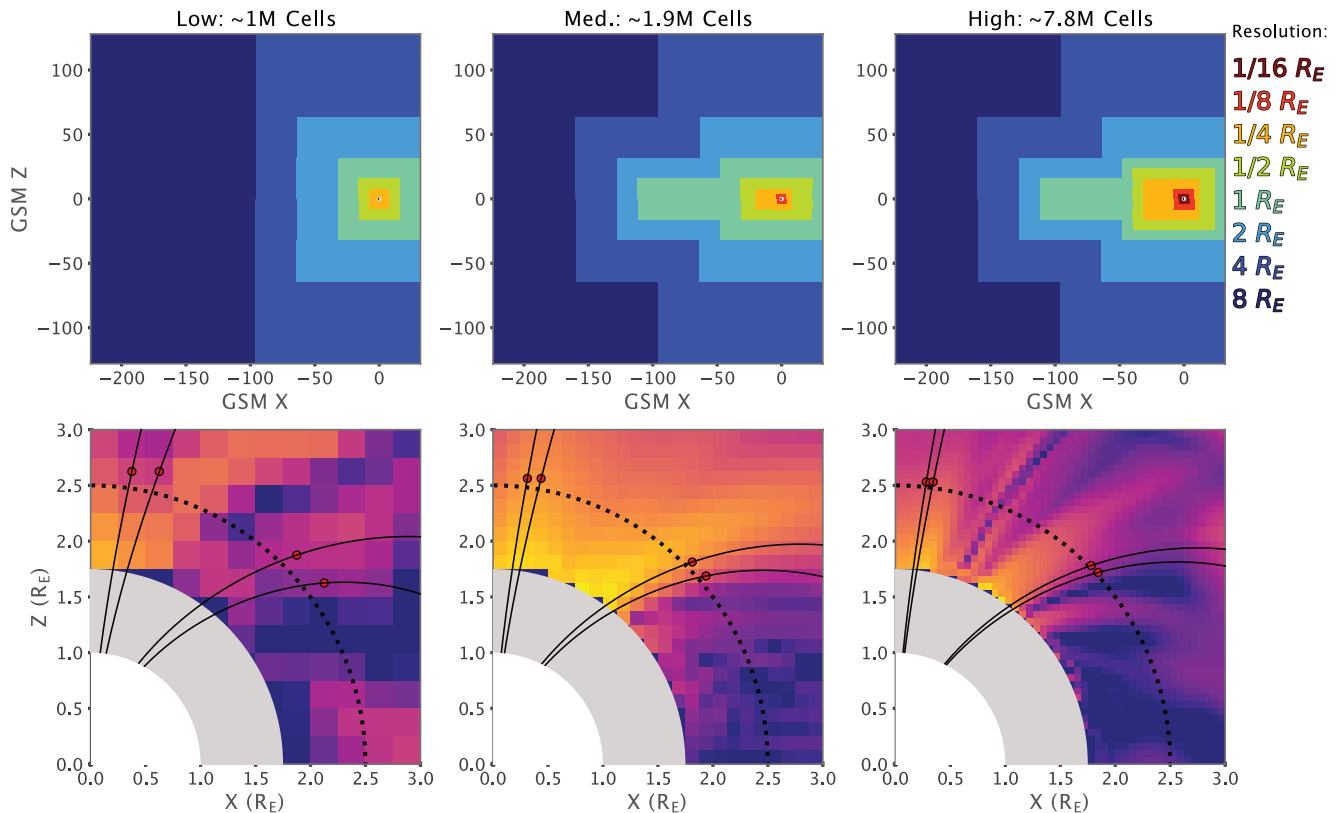
this model performs well in matching the amplitude and timing of large GICs. Since the motivation for this study is to investigate the role of different spatial resolution runs, then small deviations in these parameters would not affect the conclusions of the study.

### 2.3. SWMF Run

Simulations of the event of interest are generated using the SWMF (Tóth et al., 2005, 2012). The SWMF is a flexible software framework that executes, synchronizes, and couples otherwise independent models of different domains of the space environment. For this study, three models are employed: the BATS-R-US MHD model (D. De Zeeuw et al., 2000; Powell et al., 1999), the Ridley Ionosphere Model (RIM; Mukhopadhyay et al., 2020; Ridley et al., 2004), and the Rice Convection Model (RCM; Sazykin et al., 2002; Toffoletto et al., 2003). The models within this configuration of the SWMF are coupled to keep the aggregate solution fully self-consistent at an interval of 10 s. Coupling between RCM and the other models is described in D. L. De Zeeuw et al. (2004). RIM uses an empirical conductance model thoroughly described in Mukhopadhyay et al. (2020). Details of inter-model coupling and an overview of SWMF performance with respect to observations are reviewed in D. Welling (2019).

For this study, the configuration of the models closely follows that of Pulkkinen et al. (2013) except for the grid resolution of the BATS-R-US model. Three different grid layouts are used, illustrated in Figure 2. In the top row, the colors show the regions of different resolutions corresponding to the colored numbers to the right e.g., dark blue corresponds to regions of 8 RE point separation and dark red to 1/16th, etc. In the bottom row, the color indicates the current density, the inner white circle is the Earth, the outer gray circle is the code inner boundary of the MHD simulation. The black dotted line shows the radius at which FACs are sampled for coupling to the ionosphere solver to avoid inner boundary effects. The solid black lines are sample magnetic field lines that thread neighboring grid cells at  $R_{FAC}$ . They illustrate how the MHD grid resolution maps to spatial structure in the ionosphere/Earth's surface, which is highlighted by the red markers.

The lowest resolution uses 1 million cells with a minimum inter-cell spacing of 1/4 Earth Radii ( $R_E$ ). This configuration is the same as what is used in operations at NOAA's Space Weather Prediction Center. A



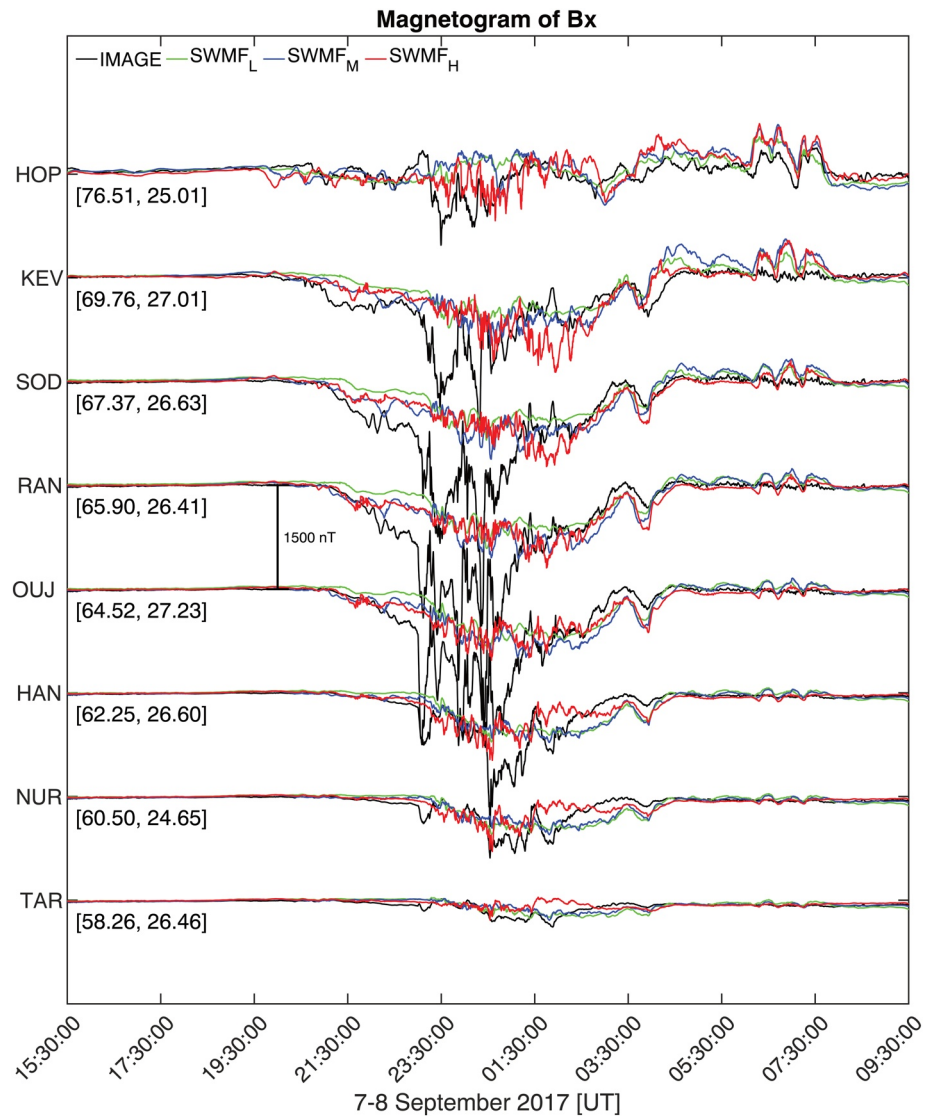
**Figure 2.** The top row shows the spatial resolution of the low medium, and high runs in the GSM meridian plane. The lower row shows the current density for each run. It can clearly be seen in the lower row that improved resolution of field-aligned currents near the inner boundary is achieved for the higher spatial resolution run.

finer resolution configuration is also used, reaching 1.9 million cells and a minimum grid spacing of  $1/8 R_E$ . Finally, a much higher resolution configuration is employed, reaching 8 million cells and a minimum spacing of  $1/16 R_E$ . The finer grid resolutions allow the code to capture smaller scale structures throughout the domain and better resolve field-aligned currents near the inner boundary.

#### 2.4. Telluric Effects

It should be noted that in reality, ground magnetic field measurements are a complex combination of the internal (telluric) and external (ionospheric) contributions. In this paper, we do not separate these and therefore some differences could arise between observations and model results. For example, larger  $dB/dt$  would be expected in some circumstances from the measurements (internal + external components) compared to the MHD run (only external component). If these are used to model GICs, then one would expect larger GICs when using the measured magnetic field as an input. It has been a common practice to neglect the contribution by induced (telluric) currents in the conducting earth when the ground magnetic field is determined from MHD simulations. This is a reasonable approximation when the amplitude of the (horizontal) magnetic field is considered. In this study the goal is to compare the results from multiple SWMF runs, so the conclusions should be unaffected. Nevertheless, it may be important for readers and future studies to mention that Juusola et al. (2020) have recently shown, using IMAGE magnetometer data, that the horizontal  $dB/dt$  is dominated by telluric currents nearly at all IMAGE stations. This means that much of the spatial and temporal structure seen in the total measured  $dB/dt$  comes from underground. For the horizontal components, ignoring the telluric effects leads to underestimations of  $dB/dt$ . Details depend on the ground conductivity, which in the IMAGE region has strong lateral gradients due to ocean coastlines and also due to inland conductivity anomalies. Since the geoelectric field is even more sensitive to conductivity, accurate GIC modeling is a big challenge. A strictly physics-based way to analyze this in full detail





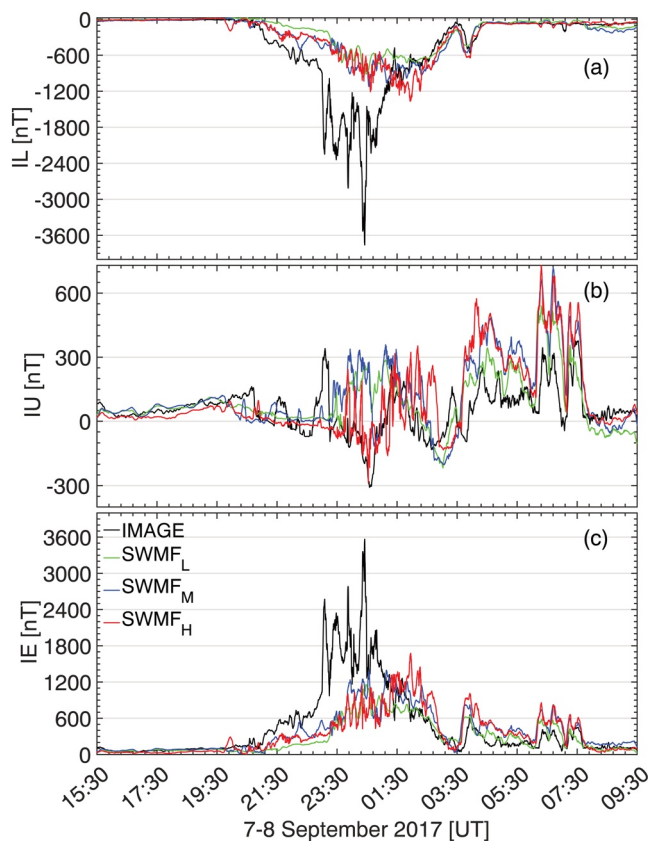
**Figure 3.** Measurements by IMAGE stations and corresponding virtual magnetometers at increasing latitudes (bottom-top) and similar longitudes. The locations of these stations are indicated by the blue markers in Figure 1. The black traces correspond to IMAGE measurements whereas the remaining lines are from the virtual magnetometers.

would be a combination of space plasma simulations and 3-D ground conductivity models. So far, very few attempts have been performed, for example, by Ivannikova et al. (2018) and Marshalko et al. (2020). The spatial variability of the ground conductivity can be substantial (e.g., Kelbert, 2020), so tailored local modeling is a necessity for obtaining accurate GIC predictions, which is not the purpose of this study.

### 3. Results

#### 3.1. SWMF-IMAGE Comparison

In this section, we compare the IMAGE magnetometer data with the corresponding virtual magnetometers from each SWMF run. Plotted in Figure 3 is a magnetogram ( $B_x$ ) from September 7, 2017 15:30 to 09:30 September 8, 2017 based on low-high latitude IMAGE stations (bottom-top). These stations are shown by the blue markers in Figure 1a and listed in Table 1 and were selected because they provide a large range of latitudes while retaining similar longitudes. The black trace corresponds to the measured data, whereas the



**Figure 4.** Comparison of the IMAGE derived geomagnetic indices IL, IU, and IE between 15:30 on September 7 and 09:30 on September 8, 2017. The IMAGE indices can be considered as local MLT equivalents to AL, AU, and AE, but with additional coverage in latitude. The black lines are the indices derived from the stations included in Figure 3. The green, blue, and red traces are from the equivalent virtual magnetometers for low, medium, and high-resolution SWMF runs, respectively. SWMF, space weather modeling framework.

green, blue, and red lines are from the SWMF low, medium, and high resolution runs, respectively. This color scheme is maintained throughout this manuscript.

As expected, there is a significant latitudinal dependency on the magnitude of  $\Delta B$  in which stations between latitudes of  $65^{\circ}$ – $70^{\circ}$  experience the largest perturbations (3,000–4,000 nT at 00:00 UT). What is striking here is that the virtual magnetometers do not capture this behavior very well, and the perturbations are significantly underestimated ( $\sim 900$  nT around 00:00 UT). Not only this, but the sharp sudden decreases are not included at all in the model, and therefore the temporal evolution of  $B_x$  is poorly captured at this stage. This is a period of strong substorm activity and MHD substorms do not create as strong of an AL signature as their real-world counterparts. A possible reason for this is the conductance model that we use (see Mukhopadhyay et al., 2020) since it is an empirical conductance pattern that tends to produce broad conductance patterns, which may allow our electrojets to smear out. The timing of the model substorm could be later than the one observed, which could account for the model  $B_x$  being delayed compared to the observations.

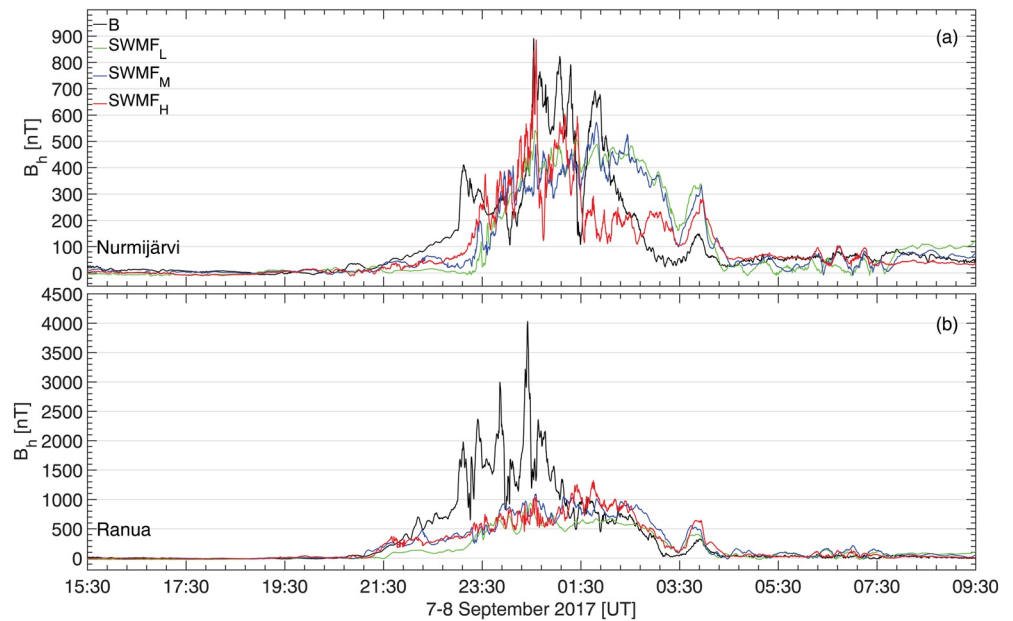
The model performance improves at around 01:15 UT in which the  $B_x$  overall trend is now captured quite well, and the differences between the model and observations are on the order of a few hundred nT. Around 03:30 UT the larger-scale variations are evident in the model which matches well with reality. The situation improves significantly at latitudes below HAN (62.25 26.60), and we see that a sharp feature around 01:15 UT is captured by the model by a sudden decrease in  $B_x$ . This feature is also observed at the neighboring stations of OIJ and NUR. In summary, some features are not captured at all by any MHD model run, whereas others are captured quite well, but this seems to be dependent on latitude and possibly the modeling and timing of substorms.

Comparing the different resolution runs, there are mixed results. In one sense the variability of the time series notably increases suggesting that smaller temporal features are better resolved. On the other hand, the magnitude of  $B_x$  is affected, which at some stations appears correct (HAN), but then diverges from the measured  $B_x$  at others (SOD). Therefore, the spatial resolution of the simulation plays a role in modifying the

large and small-scale features of  $B_x$ . What should be mentioned is that the resolution does not improve the performance between  $65^{\circ}$  and  $70^{\circ}$  around 23:30 UT, where the largest model-data differences are observed. We will investigate this more quantitatively later in the manuscript by comparing the results from NUR and RAN in much greater detail since there is a large contrast in the performance from these two stations.

We can obtain an overview of this large latitudinal behavior by computing geomagnetic indices from this collection of stations. Therefore, the IL, IU, and IE are calculated and plotted in Figure 4. These are calculated similar to AL, AU (amplitude upper), and AE (auroral electrojet), but in this case are the minimum, maximum, and range from  $B_x$  over the blue stations shown in Figure 3. The main difference is that although the IMAGE indices contain latitudinal coverage, they are restricted to local MLT. Note that, typically, IMAGE indices do not use these exact stations but all stations. For this study, we select these to support Figure 3.

As expected from Figure 3, the magnetic depression is significantly underestimated in panel (a), in this case by a factor of around three. In general, the overall trend is similar but the large-amplitude short duration spikes are not captured by the model regardless of the spatial resolution. The magnitude of the IU index in panel (b) is more representative of reality, but the higher resolution run appears to contain a larger degree of fluctuations compared to the measured IU and lower resolution runs. Panel (c) supports the other panels and clearly shows the larger fluctuations in the higher resolution run. Since these large-scale indices are

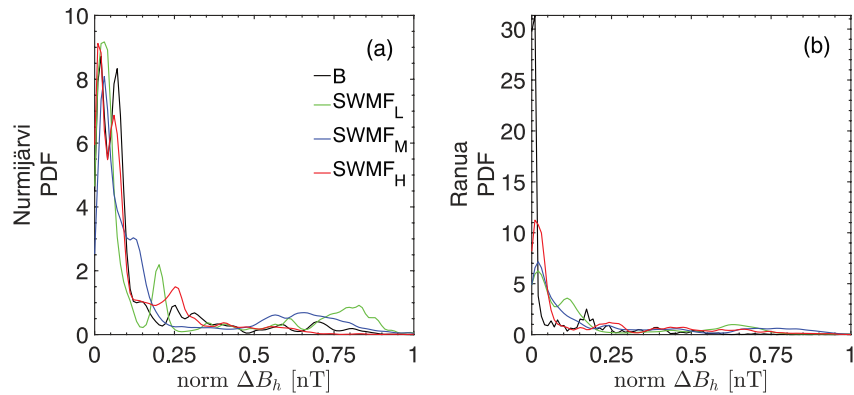


**Figure 5.** Comparison of the magnitude of the horizontal magnetic field  $B_h$  from real and virtual magnetometer data at Ranua and Nurmijärvi. The black lines correspond to the measured  $B_h$  whereas the green, blue, and red traces are  $B_h$  from the virtual magnetometers associated with the low, medium, and high-resolution SWMF runs, respectively. Note that the y-axis scaling is not consistent between the two panels. SWMF, space weather modeling framework.

dominated by the envelope limits of the magnetograms shown in Figure 3 then it highlights the difficulty of reproducing geomagnetic indices with global MHD in which the performance differs significantly as a function of latitude.

In Figure 5, we can compare the observed and modeled response in greater detail by considering the NUR and RAN stations in panels (a and b) respectively. In this case, we plot the horizontal component defined as  $B_h = \sqrt{B_x^2 + B_y^2}$ . At NUR, the peak magnitude is reproduced by the model almost exactly at 00:30 UT but then underestimates  $B_h$  afterward until around 02:45 UT. The comparison between resolutions becomes even more complex since  $\text{SWMF}_H$  underestimates the field between 01:30 and 02:45 UT whereas the other runs are more similar and then over-estimate  $B_h$ . What is significant here, is that by increasing the resolution we were able to capture higher and more true amplitudes corresponding to sharp features of  $B_h$  at NUR, although we note that some features are not captured between 01:30 and 03:30 UT. Panel (b) shows the same interval at the RAN station and there is a stark contrast to panel (a). As shown before,  $B_h$  is significantly underestimated at the beginning of the period but matches the observations well from 01:30. Although the higher resolution tends to increase the magnitude and variability of  $B_h$ , it remains substantially lower than the true values. The sharp variations of  $B_h$  likely originate from multiple substorms that are not captured by the model in this interval. This is also supported by favorable agreement after the substorm period from 01:30 UT suggesting that the global-scale underlying current systems are captured quite well, but not during strong multiple substorms. From this figure, we can easily understand why the performance of models based on data comparisons can differ substantially between different events.

In Figure 6, we plot the Probability Density Functions (PDF) of  $\Delta B_h$  after it has been normalized between 0 and 1. Or in other words,  $B_h$  is scaled by the maximum so the range of the two data sets is the same. This allows a more direct comparison between the time series which would highlight if the model is systematically underestimating the measured field. For example, if there was a systematic under-amplitude estimation, then the PDF features would be similar after being normalized. Panels (a and b) correspond to NUR and RAN stations, respectively, and the color scheme is retained from the previous figures. The results from NUR (panel a) are encouraging. There is a double peak in the observed field between 0 and 0.1 which is captured by the high-resolution run and not the lower resolutions. There are also bumps around 0.25 which are included in the red trace. Therefore, when normalized by range, the higher resolution run at NUR is



**Figure 6.** PDFs of normalized  $\Delta B_h$  for Nurmijärvi (a) and Ranua (b). The PDFs are computed based on a kernel density estimation. The interval in which each PDF is computed corresponds to that shown in Figure 5.

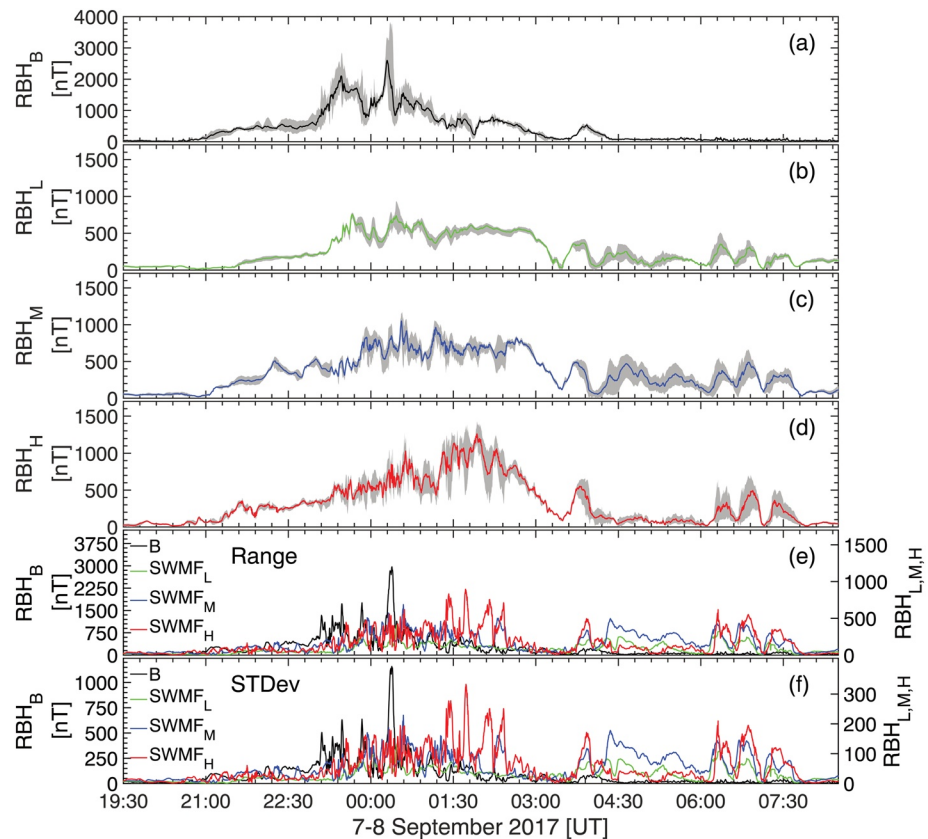
generally reproducing the distribution of  $B_h$  within the interval, which is expected considering the range of  $B_h$  in Figure 5a are similar for the measured and high-resolution time series. Interestingly, it appears the model introduces other fluctuations that are not observed in the data as seen in the low and medium-resolution runs around 0.75. We will discuss the effects of these later in the manuscript. In general, RAN measurements are dominated by several extremely large features and many small variations that explain the shape of the PDF in panel (b). The modeled RAN response shows a smaller range of variations and does not contain the large spikes seen in reality. Therefore, the overall effect is that the values of  $B_h$  are more evenly distributed. This confirms that the modeled  $B_h$  at RAN is not capturing the amplitude and behavior of the observation, and these differences are not the result of systematically under-estimating reality.

### 3.2. The Impact on Regional Structure

We now consider how the three SWMF runs differ in terms of the spatial structure of  $B_h$  over several hundred kilometers. In line with Dimmock et al. (2020), we extract a subset of IMAGE stations where the network is particularly dense and inter-station separations are on the order of around 100 km. These stations are indicated by the yellow markers in Figure 1a and listed in Table 1. Using all stations, we compute the regional average of  $B_h$  (RBH) at each time instant, which is the mean of all station measurements at each time. Then we compute the minimum and maximum of  $B_h$  at each time to provide a measure of the variability. Plotted in Figures 7a–7d are the regional average (RBH) of the magnetometers (a) and virtual magnetometers (b and c). Though all panels black corresponds to measured magnetic field whereas green, blue, and red indicate the modeled field at low, medium, and high resolutions, respectively. The grey shaded regions around each line indicate the variability determined by the maximum and minimum values over this group of stations at each time instant, i.e. large shaded areas indicate when the regional spread is high. The variability is shown more quantitatively in panels (e and f). Panel (e) shows the range of the shaded regions and panel (f) is the standard deviation measured at each time across the station set. Please note that panels (e and f) have two y-axis scales, which correspond to the measured (left) and modeled (right) time-series, and the y-axis range in panel (a) also differs from panels (b–d).

RBH follows the trend at these latitudes which appeared in Figure 3. Additionally, we see that over a few hundred kilometers there is a large variability of  $\Delta B$ . In the observed case, the largest variability occurs around 00:20 which is seen by the large shaded region in panel (a) and the high ranges and standard deviations in panels (e and f), respectively. This instance does not represent the largest regional variability in the models, which suggests the regional structure is substorm related. Interestingly, SWMF<sub>H</sub> shows a high degree of variability around 01:30 when the variability in the observed data is low compared to the earlier interval during the substorms. Nevertheless, if we consider the model-data differences in magnitudes at this time, then the two are comparable. Therefore, the comparisons of regional variabilities do converge in some limited intervals. An obvious result if we just compare the three resolutions is that SWMF<sub>H</sub> does exhibit a higher-level of spatial variability, as we expect due to the presence of small temporal features in the time-series. Another important aspect is that the regional variability between the different resolution runs is





**Figure 7.** Regional variability according to the measured and modeled ground magnetic field across the yellow marked stations shown in Figure 1 and Table 1. Panel (a) shows the measured magnetic field whereas panels (b–d) are the virtual magnetometer data at low, medium, and high resolutions, respectively. The grey shaded regions in panels (a–d) are the minimum and maximum limits across the station subset, which indicate the regional variability. Panel (e) shows the range of the shaded regions and panel (f) is the standard deviation of the magnetic field across the stations at each time step. Increases in the shaded regions and sudden enhancements in panels (e–f) indicate a large degree of regional variability.

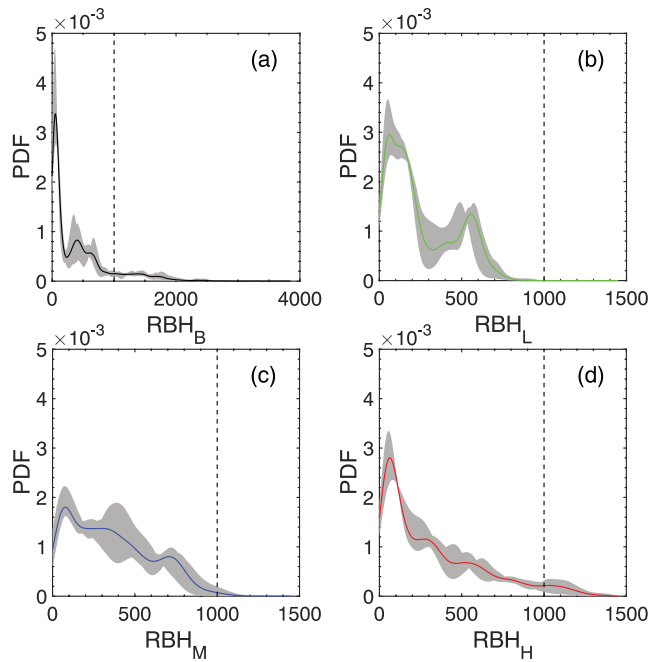
not always correlated. For example between 01:30 and 03:00,  $SWMF_H$  shows a high regional spread but the other resolutions do not. This is more obvious in the bottom two panels.

To add some context to Figure 7, the PDFs of RBH are plotted in Figure 8 where the shaded regions show the limits from the PDFs of the other stations within the region. Compared to  $SWMF_L$  and  $SWMF_M$ , the higher resolution run contains higher amplitude variations as shown by the presence of  $RBH$  above 1,000 nT (vertical dashed line). These values originate from the period between 01:30–03:00 which is why it is also accompanied by regional spread indicated by the larger gray area. In each PDF, there is a high regional spread around 200–500 nT, which also seems consistent with the observed  $RBH$ , however, this seems to be independent of the resolution and may be due to the variability of longer period features captured by all the runs. Although  $SWMF_H$  yields higher amplitudes and higher regional variability, a direct comparison with the measured  $RBH$  is difficult due to the missing high amplitude features. Nevertheless, we see more clearly here that higher resolution provides larger amplitudes and generally more spatial structure. When more stations become available at lower latitudes in the IMAGE network, this procedure should be repeated since a better agreement is expected.

### 3.3. Impact on GIC Modeling

In this section, we investigate the impacts that the model-data and model-model differences have on modeling GICs. Here we use the methods outlined in Section 2 to model GICs based on local ground conductivity





**Figure 8.** PDFs of  $RBH$  and the corresponding variability. The solid lines in each plot are the average of all the PDFs of  $BH$  from the individual subset stations whereas the gray shaded areas show the range based on the individual PDFs. Panel (a) corresponds to the PDF of  $RBH$  from the real magnetometers. The remaining panels (b–d) represent PDFs of  $RBH$  from the low, medium, and high-resolution SWMF runs. The vertical dashed line highlights a point when the high-resolution run (d) differs from the other resolutions. Note that the  $x$ -axis scaling in panel (a) differs from the remaining panels. SWMF, space weather modeling framework.

and magnetic field. For the following result, it is important to note that the SWMF virtual magnetometers lack the telluric contribution to  $B$ . It leads to underestimated  $E$  even if the simulated  $B$  (and  $dB/dt$ ) were accurate. In practice, the main purpose is to cross-compare the results using the different resolutions, and excluding the telluric contribution should not affect this comparison and the bearing of our results.

### 3.3.1. Case 1: Mäntsälä Under Measured Ionospheric Conditions

In Figure 9, the time series of  $|GIC|$  at the Finnish natural gas pipeline at Mäntsälä are compared for the interval between 20:00 UT 7 September 2017 and 05:00 UT 8 September 2017. Panel (a) is the observed GIC ( $GIC_O$ ) whereas panel (b) is GIC modeled using Nurmijärvi magnetic field and the local ground conductivity ( $GIC_B$ ). Panels (c and d) also use the local ground conductivity but the input is taken from a virtual magnetometer from the low, medium, and high spatial resolutions ( $GIC_{L,M,H}$ ), respectively. In each panel, the green, blue, and red horizontal lines are thresholds corresponding to 85%, 95%, and 97.5% quantiles (6.1, 9.7, and 12.29) of the measured GIC shown in this interval. There is no statistical significance to these quantile values but they were chosen to highlight the various amplitudes of  $|GIC|$  between the different panels. It is important to state here that in an operational setting the values of thresholds will be highly important, and therefore should be determined more analytically based on the susceptibility of the properties of the system and location. The bottom panel shows the Ratio of the Maximum Amplitudes (RMA) within 30-min intervals computed from:

$$RMA = \frac{\max(|X_{mod}|)}{\max(|X_{obs}|)}, \quad (2)$$

where  $X_{mod}$  and  $X_{obs}$  are modeled and observed signals, respectively.

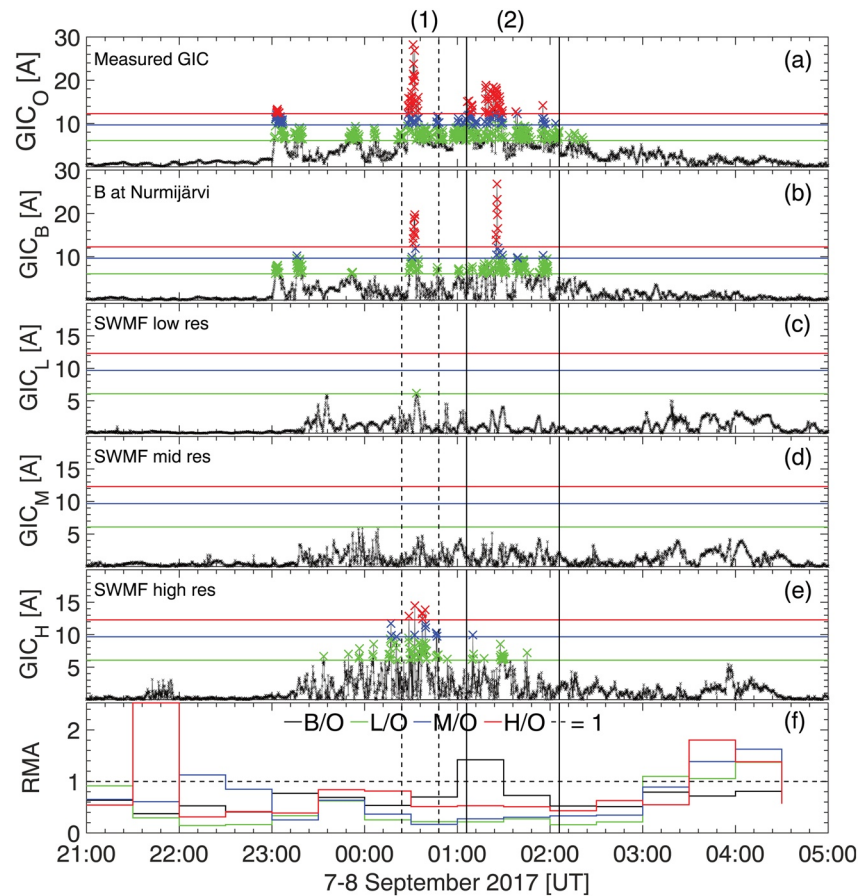
The (1) and (2) labels indicate key regions where GICs are enhanced in panel (a). We see that the observed GICs in panel (a) reaches almost 30

A, which is considered to be unusually larger for this data set. (Dimmock et al., 2019) showed that GICs for this event were in the top 10 since the recordings began.

The goal is to determine if the model will reproduce the amplitude and timing that are apparent in panel (a). As expected,  $GIC_B$  reproduces  $GIC_O$  very well in timing, amplitude, and overall temporal evolution. What becomes immediately apparent is that the resolution has a significant impact on the modeled GIC from panels (c–e). In general, the amplitude of  $GIC_H$  is higher compared to  $GIC_{L,M}$  by around 2.5 times. This is most clearly demonstrated by the fact that neither  $GIC_L$  nor  $GIC_M$  exceed the middle threshold. Although the exact amplitude is not reproduced, it is encouraging that  $GIC_H$  exceeds the 97.5% quantile of the measured GIC in this interval.

Artificial variations introduced could enhance GIC, but this is unlikely since the high amplitudes of  $GIC_H$  are in temporal agreement with  $GIC_O$  according to intervals (1) and (2). Of course, we do not see exact point-by-point instantaneous agreement, but as pointed out in previous studies (Pulkkinen et al., 2013; Yu & Ridley, 2008) this is not expected. If we adopt a more binary approach, we see that within a certain window,  $GIC_H$  exceeds a higher threshold than the low and mid resolution runs. As a result,  $GIC_H$  is a closer match to reality. This is also reflected by the RMA in panel (f) since the red line is in better agreement with the black line compared to the blue and green.

More details can be revealed in the frequency domain, therefore the Power Spectral Density (PSD) of the  $GIC_{O,B,L,M,H}$  are plotted in Figure 10a. Consistent with the previous figure,  $GIC_H$  has larger amplitudes, and the amplitude of the modeled amplitudes increases with the spatial resolution. There is one clear trend, which is that until around  $10^{-2}$  Hz, the red and black traces (both solid and dashed) agree favorably. However,

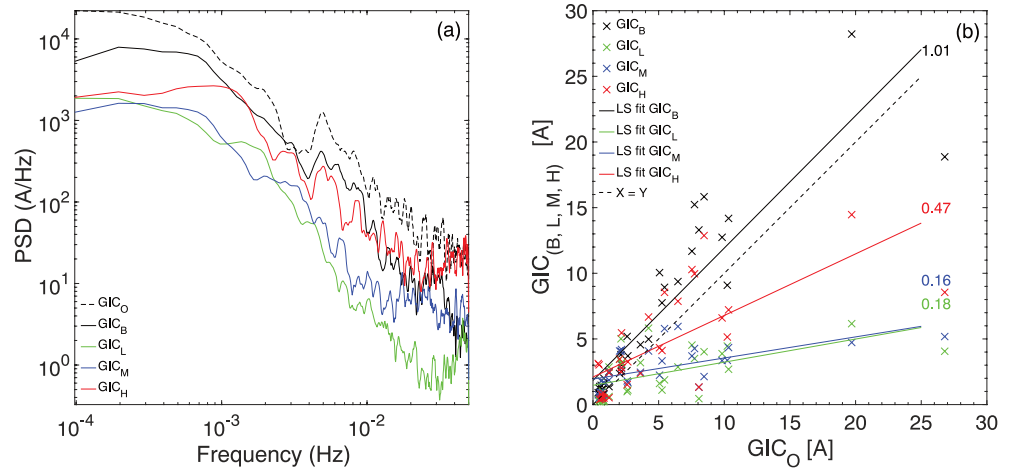


**Figure 9.** GIC in a Finnish natural gas pipeline compressor station at Mäntsälä on September 7–8, 2017. Panel (a) corresponds to the measured GIC whereas panels (b–e) are modeled from: IMAGE data, SWMF low resolution, SWMF medium resolution, and SWMF high resolution runs, respectively. Panel (f) shows the Ratio of the Maximum Amplitudes (RMA) within 30-min intervals. Note that panels (c–e) plotted on smaller y-axis scales than panels (a and b). GIC, geomagnetically induced currents; SWMF, space weather modeling framework.

there appears to be a higher power at frequencies above 10–2 Hz which are noticeable in  $GIC_H$  and  $GIC_M$ . This suggests that the model is introducing some artificial spectral components which may need to be removed. Panel (b) shows the modeled versus measured maximum GICs within a 15-min window plotted against the observed. The solid lines correspond to the least-square fits. Ideally, the slope of each line would be one, so that the modeled and the observed amplitudes are equivalent. We can see that not only does the higher resolution run give higher amplitudes, but it is also better matched to the observed GIC. Although the gradient of  $GIC_H$  is 0.74, it is notably better than  $GIC_{L,M}$  which are 0.18 and 0.16, respectively.

### 3.3.2. Case 2: Mäntsälä Under Different Ionospheric Conditions

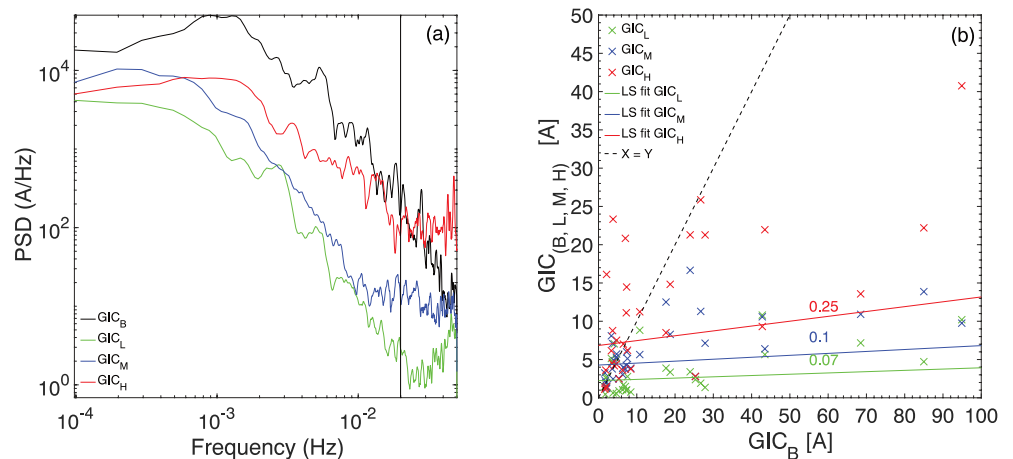
It was demonstrated in Figure 5a that the SWMF high-resolution run performed well at NUR and captured some of the complex features of the disturbances. However, according to Figure 3, this is not the case for the higher latitude stations. Figure 5b showed that for Ranua the model performed quite poorly and underestimated the geomagnetic disturbance by around a factor of three. Here we now assess the impact this would have on modeling GICs. As an exercise, we can investigate what would have been the outcome if this disturbance occurred at Mäntsälä? GIC recordings for this hypothetical scenario do not exist, but as shown in Figure 9, and other studies (Dimmock et al., 2019), the GIC at Mäntsälä can be modeled reasonably accurately using magnetic field measurements and the local ground conductivity. Therefore, we take the magnetic field (observed and virtual) at Ranua, and apply this to the Mäntsälä ground conductivity and GIC model. In effect, we are moving the ionospheric currents and not the pipeline.



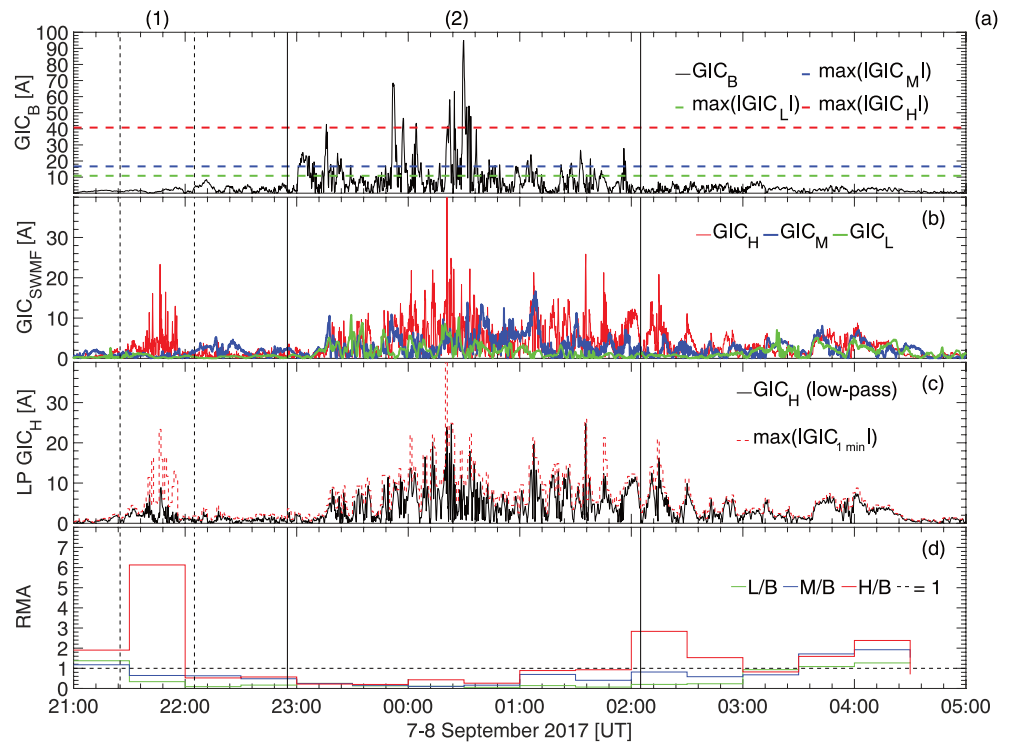
**Figure 10.** Panel (a) shows the power spectral density of the GIC time series shown in Figure 9, which was smoothed by 10 data points to aid visual inspection. Plotted in panel (b) are the modeled GIC from the real and virtual magnetometers against the measured GIC. In each axis, the value is the absolute maximum within a 15-min window covering the interval shown in Figure 9. The solid lines are least-square fits and the dashed black trace indicates when the x and y values are equal. GIC, geomagnetically induced currents.

We begin with Figure 11 by reproducing Figure 10, with the difference that  $GIC_B$  is adopted the reference rather than  $GIC_O$ . Panels (a) and (b) show  $GIC_B$  and the modeled GIC for each resolution, respectively. The maximum amplitude from each modeled GIC is shown in panel (b) by the horizontal dashed lines. The black trace in panel (c) is the low pass filtered  $GIC_H$  whereas the dashed red line is the 1 min maximum of  $GIC_H$ . Note that the reduction in  $GIC_H$  in the low pass filtered signal before 22:00. We have also plotted the RMA similar to the previous GIC comparison to highlight how well the models are capturing the expected GICs amplitudes in 30-min intervals.

Although  $GIC_H$  demonstrates higher amplitudes, they appear consistently lower than  $GIC_B$  except at the frequencies above the solid black line. Again, it seems that artificial higher frequencies are introduced above  $10^{-2}$  Hz and the effects from this will be shown later. Panel (b) also shows that the slopes of the least square fits are considerably lower, suggesting that the models are not reproducing the amplitudes of  $GIC_B$ . The differences between the runs are also not as significant here suggesting that during cases where performance is poor, the resolution is insignificant.



**Figure 11.** This figure is equivalent to Figure 10 but the x-axis is  $GIC_B$  instead of  $GIC_O$ . The vertical black line in panel (a) indicates a deviation in PSD from the observed to the modeled. A similar feature was noticed in Figure 10a. GIC, geomagnetically induced currents; PSD, power spectral density.



**Figure 12.** Modeled GIC in the Finnish natural gas pipeline compressor station at Mäntsälä. Instead of using the local measurements, the magnetic field data from the Ranua station was used. Panel (a) is the modeled GIC from the actual RAN magnetometer. The horizontal dashed lines in panel (a) are the maximum of the data plotted in panel (b), which are the modeled GIC using the virtual magnetometers from the low, mid, and high resolution runs. Panel (c) shows the low-pass filtered (by 0.01 Hz)  $GIC_H$  data (black) and a 1-min maximum  $GIC_H$  for comparison (red). Panel (d) shows the RMA similar to Figure 9f, but the  $GIC_B$  is used as the reference as opposed to  $GIC_O$ . Two important features are the high values of RMA for the high-resolution run just before 22:00 UT and the low values for all resolutions around 00:00 UT. Regions of interest are indicated by the vertical lines and the (1) and (2) labels above the plot. GIC, geomagnetically induced currents; RMA, ratio of the maximum amplitudes.

Returning to the time domain, we plot  $GIC_{B,L,M,H}$  in Figure 12 and the RMA. During interval (2) marked at the top of the plot, it is clear that the GICs expected from these conditions ( $>90A$ ) are around three times higher compared to when the in situ conditions were used. This is reflected in the low values of RMA in panel (d) indicating an under-estimation of the ground magnetic field variations. The large spikes in  $GIC_B$  show good temporal agreement with the spikes in  $B_h$  seen in Figure 5 suggesting that substorms would contribute to driving these large GICs. This also offers an explanation why  $GIC_H$  is much lower than  $GIC_B$ . The horizontal lines in panel (a) show that even the high-resolution run is around 50A below the expected GIC and the low and mid resolutions are even lower ( $<20A$ ). Although  $GIC_H$  is enhanced during interval (2), the temporal agreement is poor if we compare to Figure 9, again suggesting that the substorms occurring during this interval are highly important to driving GICs and will dictate the amplitude and timing of high GIC events.

Interval (1) shows an increase in  $GIC_H$  (panel b), which is not realized in  $GIC_B$ , suggesting this is artificial. These variations can be removed using a low-pass filter with minimal effects on the remaining time series, however, issues may arise if frequencies overlap between artificial and physical contributions to the modeled ground magnetic field. The artificial frequencies manifest as high values of RMA in panel (d). The amplitude of these variations is small according to Figure 5b around 21:30 UT. Even though these artificial features are low in amplitude, the higher spectral content (large  $dB/dt$ ) creates large GICs during the modeling process. This would cause problems in an operational setting because they drive large amplitude GICs during this interval. It will be critical to remove or avoid these features in future studies if higher resolution runs are to be adopted and coupled to ground conductivity models. We low-pass filtered  $SWMF_H$  below the vertical line shown in Figure 12 and re-modeled the GIC, which is plotted in Figure 12c. We can see that the

artificial GICs have been significantly reduced. In practice, this demonstrates that for GIC modeling purposes, even small-amplitude variations have to be carefully considered since they can sometimes translate to high  $dB/dt$ .

#### 4. Discussion

GICs have been shown to result in physical damage to technological systems (Anderson et al., 1974; Bolduc, 2002; Boteler & Jansen Van Beek, 1999; Pulkkinen et al., 2005; Rosenqvist et al., 2005), and therefore predicting these events with models to sufficient accuracy is highly important. The objective of this study has been to assess the impact of the spatial resolution of the SWMF with regards to capturing global-to regional-scale geomagnetic disturbances, as well as the effects on modeling GICs. Recent experimental studies have highlighted the regional nature of the geoelectric field (Bedrosian & Love, 2015; Dimmock et al., 2020; Love et al., 2019; Ngwira et al., 2015, 2018; Pulkkinen et al., 2003; Viljanen & Pirjola, 2017), which has renewed the motivation to determine to what extent global MHD models can capture this behavior. We have focused on a period during the September 2017 space weather event (Clilverd et al., 2018; Dimmock et al., 2019) in which we ran the SWMF at low, medium, and high spatial resolutions (see Figure 2). Using Biot-Savart integrals we have derived so-called virtual magnetometers to compare modeled geomagnetic perturbations with those measured by the IMAGE magnetometer network. We have analyzed the differences between observed ground magnetometer data and virtual magnetometers derived from three different SWMF runs over both a large latitudinal range and a regional-scale of only several hundred kilometers. Also, using the (virtual and real) surface magnetic field and local ground conductivity, we have modeled GICs in the Finnish natural gas pipeline compressor station at Mäntsälä for all three SWMF runs. These modeled GICs were compared with the measured GICs to determine if the spatial resolution plays a role in modeling GICs.

##### 4.1. Latitudinal and Regional Variability

To determine if substorms play a large role in the model-data discrepancies, we identified Substorm Onsets and Phases From Indices of the Electrojet (SOPHIE; Forsyth et al., 2015; using the IMAGE auroral indices) to determine SOPHIE onsets that were located in the Fennoscandian sector. Note that since we used only IMAGE indices we are restricted to identifications during night-time MLT sectors. Nevertheless, this coincides with the interval that we are investigating. In the Appendix, Tables A1–A3 show the results from this between 20:00 and 02:00 UT September 7–8, 2017. According to this, multiple substorms occurred during this interval, which is consistent with the large depressions of  $B_x$  (negative bays) seen in Figure 3. Thus, we conclude that the large data-model discrepancies between geomagnetic latitudes between  $62^\circ$  and  $70^\circ$  are due to multiple strong substorms occurring over this period, which the model does not capture. Additional evidence of this is provided by the lack of sharp gradients in the synthetic IMAGE indices plotted in Figure 4. Recently Haiducek et al. (2020) reported that the SWMF has weak (but statistically significant) predictive skill for substorms. Therefore, during intervals of strong and frequent substorms, the ability of the SWMF to predict the surface magnetic field will be limited. According to our results, increasing the spatial resolution has little effect during these conditions. This could explain the disparity in the various data-model comparisons across multiple events, model configurations, and validation metrics, which has been reported in previous studies (Pulkkinen et al., 2010).

At lower latitudes ( $<61^\circ$ ) we noticed a significant improvement in the virtual-real magnetometer comparison (see Figure 5). Kwagala et al. (2020) reported that the northward ground  $\Delta B$  from SWMF was better predicted at sub-auroral latitudes ( $58^\circ$ – $67^\circ$ ) which generally agrees with our observations. In our study, one reason for this is that NUR did not measure the large  $B_x$  depressions that dominated at higher latitudes. However, there is a sharp feature around 00:45 UT (see Figure 5) that is captured by the model in both magnitude and timing. Since it is very clear from the higher-latitude model-data comparison that the SWMF is not capturing substorms at this time, then it is our interpretation that this is a non-substorm feature. A possible explanation for this could be the increase in solar dynamic pressure shown by Dimmock et al. (2020) (Figures 4a and 4b) at this time, suggesting this could be a directly driven response in which the MHD model can capture. Therefore, it is important to reiterate that it is clear that higher spatial resolution can improve accuracy - as long as we can capture, and resolve, the appropriate underlying physical process.



What is also interesting to note is that we negligible improvement from low-medium spatial resolutions (1–1.9 M cell) and moderately increasing the spatial resolution does not ensure increased performance, as one would expect, thus there is a complex dependency. Thus, this may imply that for GIC purposes, increased performance is only achieved if the resolution is increased significantly than what is typically used in operational settings.

Accurately capturing regional variability with global MHD models will remain difficult regardless of the spatial resolution. Similar to previous studies (Pulkkinen et al., 2015; Dimmock et al., 2020), spatial properties of disturbances can be investigated using a subset of the IMAGE stations where the coverage is sufficiently dense. From this procedure, higher resolution runs produced a wider range of values across the virtual magnetometers that we compared (see Figure 7). A direct comparison with the real magnetometers was difficult in this case because of the presence of strong substorms, which were not captured by the model. Substorms have been linked to regional variability (Clilverd et al., 2018; Ngwira et al., 2015, 2018; Pulkkinen et al., 2003, 2015; Viljanen, Pulkkinen, et al., 2006; Viljanen, Tanskanen, et al., 2006), and as a result, the regional variations in the virtual magnetometers should originate from different physical mechanisms. Figure 2 shows clearly that the field-aligned currents are significantly better resolved in the high-resolution run. Besides, spatial maps of ionospheric currents (not shown) imply that higher resolutions tend to exhibit stronger and smaller-scale ionospheric current patterns. This is consistent with Sorathia et al. (2020) who report that the high-resolution global MHD simulation GAMERA produced localized field-aligned current structures. Nevertheless, many observational studies have shown highly-complex ionospheric current features (Apatenkov et al., 2020) such as vortexes (Belakhovsky et al., 2019; Viljanen et al., 2006; Viljanen, Tanskanen, et al., 2006), which pose significant challenges for modelers. Especially since it is now becoming clear that models need to capture some of these features if they are to yield accurate regional predictions of geomagnetic disturbances.

Pulkkinen et al. (2010) suggested that high-latitude ionospheric currents are amplified and irregular during more active conditions due to processes such as substorms, decreasing predictability. In support of this, Engebretson et al. (2019) showed that night-time magnetic perturbation events may be connected to localized poleward boundary expansions and auroral streamers, which in turn can be linked to tail dynamics such as dipolarization fronts. Thus, although geomagnetic disturbances may manifest locally, understanding the underlying physical processes and accurately modeling them is a complex global issue. The increased regional variability reported at these latitudes (65°–70°) in the model is likely associated with non-substorm processes such as magnetospheric compressions, sudden storm commencements, and increased magnetospheric convection, and are likely captured better with the higher resolution run in this study. Although the model was not able to reproduce the measured regional variations, the fact remains that capturing non-substorm regional variability will be key to improving the accuracy of lower latitude forecasts where substorms may not be as dominant. It is also worth highlighting that many large population centers occupy lower latitudes in Fennoscandia. For example, the geographic latitude of the NUR station is similar to Helsinki, and Stockholm. These latitudes are of particular interest since southern Sweden has experienced power disruptions due to GICs (Wik et al., 2009) in the past; which is suggested to be due to the lower crustal conductance (Rosenqvist & Hall, 2019). This makes the area prone to high amplitude geoelectric fields and disruptions due to GICs. Recent studies such as Viljanen and Pirjola (2017) and Dimmock et al. (2020) have presented strong quantitative evidence that the regional variability of geomagnetic disturbances is important when modeling GICs.

#### 4.2. The Impact on GIC Modeling

If we can accurately model the temporal and spatial characteristics of the surface magnetic field, then useful modeling of GICs should be a straightforward task. This is dependent on knowing the physical properties of the system in question, and possessing a realistic knowledge of the local ground conductivity (Boteler, 2003; Bedrosian & Love, 2015; Pulkkinen et al., 2017; Kelbert & Lucas, 2020). At Mäntsälä, the conductivity is provided by Viljanen, Pulkkinen, et al. (2006); Viljanen, Tanskanen, et al. (2006), and as shown by Equation 1, there is an exact relationship between the geoelectric field and the and GIC (Pulkkinen, Pirjola et al., 2001; Pulkkinen, Viljanen et al. 2001) when the geoelectric field is uniform. In the case of nonuniform geoelectric fields as in this study, a local value can be used as an approximation that has been shown to perform well.

We first modeled GIC using the measured and modeled magnetic field at NUR, and the measured data was in excellent agreement with the measured GIC (see Figure 9) as expected (Viljanen, Pulkkinen, et al., 2006; Viljanen, Tanskanen, et al., 2006). Regarding the modeled magnetic field, the high-resolution run stood apart from the mid and low resolution runs; as it was the only time series to exceed the upper and middle thresholds at similar times to the measured GICs. Therefore, for this example, realistic amplitudes of  $|GIC|$  are only realized when GICs were modeled with the high-resolution SWMF. In fact,  $\max|GIC_H|/\max|GIC_M| = 2.43$  and  $\max|GIC_H|/\max|GIC_L| = 2.35$ , showing that the peak GICs in the higher resolution run were almost 2.5 times that in the mid and low resolution runs. Also, the  $GIC_H$  exceeded the 97.5% quantile of the measured GIC. The Heidke skill scores (HSS) shown in Table B2 show an increase in skill (from low-high) of 0.16–0.74, 0–0.56, and 0–0.48 for the lower, middle, and upper thresholds, respectively. The improvement for the higher resolution result is explained by the fact that the high-resolution run generally exhibited enhanced power spectral density across the signal spectral range (see Figure 10a). Physically, this means that the higher resolution run is capturing increased complexity in the behavior of ionospheric currents, which then manifests as a more accurate representation of the surface magnetic field across the available frequency domain. There was also a higher correlation (line of the best fit coefficient of 0.47) between the peak values (see Figure 10b) suggesting that the model agreed well across the interval we tested at NUR. Since substorms did not seem to dominate the time series, this result implies that higher spatial resolution may significantly improve the capability of global MHD models to forecast GICs during similar situations. Keeping in mind that we used a metric of exceeding a given threshold within a specified window, which has been adopted by other studies (Tóth et al., 2014; D. T.; Welling et al., 2017). Thus one key result here is that higher spatial resolution provided an increase in performance, as long as we resolved the underlying feature.

We used the magnetic field at RAN to model the GICs at Mäntsälä, effectively shifting the ionospheric response at RAN to the lower latitude of Mäntsälä. Since measured GICs do not exist for this hypothetical scenario, we used the modeled GIC from the RAN magnetometer as the reference/ground truth. RAN observed multiple large ( $\sim 2000 - 4,000$  nT) substorms, which occurred throughout two hours. What was striking here is the significantly increased GICs of around 100 A. Compared to NUR, this is a three times increase, re-enforcing the importance of substorms in large GIC events. We should note that  $\max|GIC_H|/\max|GIC_M| = 2.45$  and  $\max|GIC_H|/\max|GIC_L| = 3.77$ , so the high resolution run does provide higher amplitude GICs compared to the lower resolutions. Interestingly the  $GIC_H$  exceeds a 97.5% quantile ( $\sim 30$  A) and the low and mid-resolution runs did not. This is reflected in the 0.33 HSS for the upper threshold compared to 0 for the other runs, suggesting even in this case, with substorms, the higher resolution run performed better. Nevertheless, there is still an underestimation of the peak GICs of around 50 A. The most important point, as shown in Figures 11 and 12, is that the temporal agreement between the maximums is poor. This remains true across the entire interval since the relationship (slopes of least square fits in Figure 11) between the peak GICs were comparable across all the model resolutions. This is in stark contrast to the previous GIC comparison and re-iterates that if substorms are dominant and frequent, then unless a model can capture these events to some reasonable degree, the spatial resolution will not offer a reasonable solution. As pointed out by Pulkkinen et al. (2010); Pulkkinen et al. (2011), assessing model performance is heavily dependent on the stations, events, and validation metrics that are adopted.

An unexpected result of this study was that the model occasionally introduced artificial frequencies to the magnetic field time series which correspond to periods above  $\sim 50$  s. These appear as low amplitude ( $\sim 10$  nT) fluctuations as shown in Figure 3 at RAN around 09:30–09:50 UT. However, it is the  $dB/dt$  of the field and not necessarily the amplitude which drives large GICs, as pointed out by Dimmock et al. (2019). As a result, the GICs resulting from these artificial features were large ( $\sim 22$  A), and from an operational perspective would introduce false-positive events. This was checked by applying a low-pass filter with a cut-off of  $10^{-2}$  Hz, and then re-modeling the GICs (see Figure 12b). Although in this case, it was straightforward to eliminate these effects, problems may arise if artificial frequencies overlap with natural ones. To determine the cause of this, we re-ran the model again for this hour (21:00–22:00 7/09/2017) with a different SWMF coupling frequency and output frequency. These features remained and therefore the origin of these features is still under investigation and will be addressed in future studies. Diagnosing these issues beyond the scope of this paper, but highlights the challenges of modeling GIC with global MHD and the accuracy required in the surface magnetic field predictions.

### 4.3. Future Implications and Outlook

The work presented in this manuscript has several key implications. One obvious point throughout this study is that a significant improvement in predicting and capturing substorms is required to predict the surface magnetic field at high latitudes during strong events. If these explosive magnetotail events are not included, then increasing the spatial resolution provides a negligible improvement, and any prediction will be of low skill and non-actionable. For lower latitudes, or when substorms are not large and frequent, then a higher spatial resolution may provide a significant improvement when predicting GICs. Although this is true of this event, we have studied only one case, and it will be crucial to verify this with additional studies of more events. It is important to keep in mind that for operational versions of global MHD models, a high priority is placed on speed (faster than real-time) and robustness, and increasing the spatial resolution may not be straightforward due to the increased computational expense. Therefore, it is recommended that one should take into consideration the latitude where predictions are to be made before investing time and resources into higher resolution runs. Nevertheless, we have shown that given the correct conditions, it is feasible to model GICs with a global MHD model, ground conductivity, and an accurate description of the system. In our case, once the surface magnetic field has been obtained, GICs can be evaluated almost instantaneously.

There is still significant work to be done to achieve accurate forecasting capability from global MHD models such as SWMF. Mainly, we need to better understand the underlying driving mechanisms of complex geomagnetic disturbances and improve our ability to model these complex physical processes. For context, we have presented a single case study, and it will be important that future validation and benchmarking studies that utilize global MHD models also consider the effects of spatial resolution to shed more light on the issues raised here. Global magnetospheric plasma simulations are evolving rapidly, and future developments such as improving the ionospheric conductance (Mukhopadhyay et al., 2020), and improving tail reconnection physics by embedded PIC boxes (Daldorff et al., 2014) may improve the results at higher latitudes. These activities performed in this study should continue when updated versions of these models become available. Although the continued development of global magnetospheric models is required, this should also be achieved in parallel with advancing experimental capabilities. With improved magnetometer coverage, additional magnetotelluric surveys, and novel magnetospheric measurements, this fundamental research should also translate to improved operational capabilities. Finally, for accurate GIC modeling, the role of ground conductivity cannot be overlooked, and it is highly important to couple geomagnetic predictions from physics- and empirical-based models to realistic ground conductivity models with sufficient spatial resolution to resolve key geological features.

## 5. Conclusions

The results from this study have led to many novel insights into the role of global MHD spatial resolution on capturing large and regional scale variability as well as the impact on modeling GICs. The main conclusions from this paper are listed below.

Data-Model comparison:

1. At low to mid-latitudes within the IMAGE network, where substorm effects are not as dominant, higher spatial resolution may provide more accurate surface magnetic fields.
2. Substorms remain a significant challenge for accurately reproducing surface magnetic fields at high latitudes and will limit the accuracy of virtual magnetometers from global MHD.

Spatial resolution:

1. It is promising that a higher spatial resolution produced noticeably improved GICs in terms of amplitude and timing at NUR.
2. For the GIC modeling case at NUR, only the high-resolution run provided improved results suggesting a complex relationship between spatial resolution and performance.
3. In general, the amplitudes of horizontal geomagnetic perturbations were larger in the high-resolution run.

4. By increasing the spatial resolution, a higher degree of regional ( $\sim 500$  km) variability of geomagnetic perturbations was observed by virtual magnetometers.
5. The spatial resolution provides increased performance only if the underlying physical process is captured/resolved.

Operational consequences:

1. Models can introduce artificial spectral content to virtual magnetometers. Although these may only be a few nT, their higher frequency (i.e., high  $dB/dt$  peaks) can drive large amplitude GICs and introduce false-positive predictions. Methods to remove these effects should be considered.
2. The operational version of global MHD models adopt a spatial resolution comparable to the low resolution run here. According to Figure 9, a higher resolution may be needed in some situations to capture smaller spatial and temporal scale features important to GICs.
3. Increasing the spatial resolution may not necessarily lead to significantly improved predictive skill. Forecasters considering such a transition should consider the latitudes at which predictions are targeted and the expected magnetospheric processes which would be dominant during strong-extreme events.

## Appendix A: Substorm Identification Based on SOPHIE and IMAGE Indices

We use the Substorm Onsets and Phases from Indices of the Electrojet (SOPHIE) methodology to identify the onsets and phases of substorms which are listed in Tables A1–A3 below. The input to the technique is the IL index and the SMU index. Note that substorms will only be detected on the nightside due to the fixed positions of IMAGE in MLT. Each table lists the date, time, phase, and threshold for each identification. The phase is numbered according to 1 for growth/energy input, 2 for expansion, and 3 for recovery. The threshold refers to the percentile in the rate of change of IL index. For a complete description of this technique we refer the readers to Forsyth et al. (2015).

**Table A1**

*Times of Substorm Onsets and Phases Based on the SOPHIE Method With Threshold = 50*

Date	Time (UT)	Phase	Threshold
2017/09/07	20:11:00	3	50
2017/09/07	20:29:00	2	50
2017/09/07	22:04:00	3	50
2017/09/07	22:15:00	2	50
2017/09/07	22:29:00	3	50
2017/09/07	22:45:00	2	50
2017/09/07	23:29:00	3	50
2017/09/08	00:03:00	2	50
2017/09/08	00:22:00	3	50
2017/09/08	01:26:00	2	50
2017/09/08	01:45:00	3	50

**Table A2**

*Times of Substorm Onsets and Phases Based on the SOPHIE Method With Threshold = 75*

Date	Time (UT)	Phase	Threshold
2017/09/07	20:15:00	3	75
2017/09/07	20:30:00	2	75
2017/09/07	22:25:00	3	75
2017/09/07	22:49:00	2	75

**Table A2**  
*Continued*

Date	Time (UT)	Phase	Threshold
2017/09/07	23:28:00	3	75
2017/09/08	00:03:00	2	75
2017/09/08	00:22:00	3	75
2017/09/08	01:26:00	2	75
2017/09/08	01:45:00	3	75

**Table A3**

*Times of Substorm Onsets and Phases Based on the SOPHIE Method With Threshold = 90*

Date	Time (UT)	Phase	Threshold
2017/09/07	20:34:00	2	90
2017/09/07	22:27:00	3	90
2017/09/07	22:52:00	2	90
2017/09/07	23:28:00	3	90
2017/09/08	00:03:00	2	90
2017/09/08	00:22:00	3	90
2017/09/08	01:26:00	2	90
2017/09/08	01:44:00	3	90

## Appendix B: Heidke Skill Scores for GIC Modelling

Here calculate the Heidke Skill Scores (HSS) corresponding to Figures 9 and 12. The HSS is computed according to

$$HSS = \frac{2(AD - BC)}{[(A + C)(C + D) + (A + B)(B + D)]}, \quad (B1)$$

**Table B1**

*Contingency Table Used to Calculate HSS*

Event forecast	Event observed	Event not observed
Yes	A	B
No	C	D

where A, B, C, and D are determined from a contingency Table B1.

Table B2 shows the HSS values based on exceeding three specific thresholds. The three thresholds were based on the 85%, 95%, and 97.5% quantiles of the reference GIC. For the NUR and RAN cases, these are  $GIC_O$  and  $GIC_B$ , respectively. It is important to mention here that these thresholds are only used to measure the skill for this event and are only applicable to this case study. In general, these threshold values should be more rigorously evaluated based on statistical analysis of previous events.



**Table B2**  
*Heidke Skill Scores Corresponding to 9 and 12*

Data set	Threshold [nT]	Input	HSS
GIC <sub>B</sub>	6.1	NUR	0.87
GIC <sub>L</sub>	6.1	NUR	0.16
GIC <sub>M</sub>	6.1	NUR	0.00
GIC <sub>H</sub>	6.1	NUR	0.74
GIC <sub>B</sub>	9.7	NUR	0.71
GIC <sub>L</sub>	9.7	NUR	0.00
GIC <sub>M</sub>	9.7	NUR	0.00
GIC <sub>H</sub>	9.7	NUR	0.56
GIC <sub>B</sub>	12.3	NUR	0.48
GIC <sub>L</sub>	12.3	NUR	0.00
GIC <sub>M</sub>	12.3	NUR	0.00
GIC <sub>H</sub>	12.3	NUR	0.48
GIC <sub>L</sub>	9.7	RAN	0.39
GIC <sub>M</sub>	9.7	RAN	1.00
GIC <sub>H</sub>	9.7	RAN	0.50
GIC <sub>L</sub>	19.6	RAN	0.00
GIC <sub>M</sub>	19.6	RAN	0.00
GIC <sub>H</sub>	19.6	RAN	0.47
GIC <sub>L</sub>	30.3	RAN	0.00
GIC <sub>M</sub>	30.3	RAN	0.00
GIC <sub>H</sub>	30.3	RAN	0.33

*Note.* The Thresholds are Based on 85%, 95%, 97.5% Quantiles of the Reference

## Acknowledgments

A. P. Dimmock, E. Yordanova, and L. Rosenqvist were funded by the Swedish Civil Contingencies Agency Grant 2016-2102. The work by A. Viljanen was supported by the Academy of Finland grant no. 314670. R. J. Boynton has received financial support from the UK NERC under grant NE/P017061/1. D. T. Welling was supported by the National Science Foundation award ICER-1663770. I. J. Rae and C. Forsyth are supported in part by Natural Environment Research Council (NERC) grants NE/P017150/1 and NE/V002724/1. C. Forsyth is supported by a NERC Independent Research Fellowship NE/N014480/1. We thank the institutes that maintain the IMAGE Magnetometer Array: Tromsø Geophysical Observatory of UiT the Arctic University of Norway (Norway), Finnish Meteorological Institute (Finland), Institute of Geophysics Polish Academy of Sciences (Poland), GFZ German Research Centre for Geosciences (Germany), Geological Survey of Sweden (Sweden), Swedish Institute of Space Physics (Sweden), Sodankylä Geophysical Observatory of the University of Oulu (Finland), and Polar Geophysical Institute (Russia).

## Data Availability Statement

The IMAGE data and its derived data products can be obtained free of charge at <http://space.fmi.fi/image>. The GIC data can be viewed at the following address <http://space.fmi.fi/gic/> and downloaded from [http://space.fmi.fi/gic/man\\_ascii/man.php](http://space.fmi.fi/gic/man_ascii/man.php). The space weather modeling framework and all included models used here is openly available via <https://github.com/MSTEM-QUDA>. All data from the model used for this study are available (virtual magnetometer output files and SWMF PRAM files) both as supplementary data and deposited in Zenodo with <http://doi.org/10.5281/zenodo.4608556>.

## References

- Akasofu, S.-I. (1964). The development of the auroral substorm. *Planetary and Space Science*, 12(4), 273–282. [https://doi.org/10.1016/0032-0633\(64\)90151-5](https://doi.org/10.1016/0032-0633(64)90151-5)
- Anderson, C. W., Lanzerotti, L. J., & MacLennan, C. G. (1974). Outage of the L4 System and the Geomagnetic Disturbances of 4 August 1972. *Bell System Technical Journal*, 53(9), 1817–1837. <https://doi.org/10.1002/j.1538-7305.1974.tb02817.x>
- Apatenkov, S. V., Pilipenko, V. A., Gordeev, E. I., Viljanen, A., Juusola, L., Belakhovsky, V. B., et al. (2020). Auroral omega bands are a significant cause of large geomagnetically induced currents. *Geophysical Research Letters*, 47(6), e2019GL086677. <https://doi.org/10.1029/2019GL086677>
- Bedrosian, P. A., & Love, J. J. (2015, 12). Mapping geoelectric fields during magnetic storms: Synthetic analysis of empirical United States impedances. *Geophysical Research Letters*, 42(23), 10160–10170. <https://doi.org/10.1002/2015GL066636>
- Belakhovsky, V., Pilipenko, V., Engebretson, M., Sakharov, Y., & Selivanov, V. (2019). Impulsive disturbances of the geomagnetic field as a cause of induced currents of electric power lines. *Journal of Space Weather Space and Climate*, 9, A18. <https://doi.org/10.1051/swsc/2019015>
- Bolduc, L. (2002). GIC observations and studies in the Hydro-Québec power system. *Journal of Atmospheric and Solar-Terrestrial Physics*, 64(16), 1793–1802. [https://doi.org/10.1016/S1364-6826\(02\)00128-1](https://doi.org/10.1016/S1364-6826(02)00128-1)
- Boteler, D. H. (2003). Geomagnetic hazards to conducting networks. *Natural Hazards*, 28(2–3), 537–561. <https://doi.org/10.1023/A:1022902713136>

- Boteler, D. H., & van Beek, G. J. (1999). August 4, 1972 revisited: A new look at the geomagnetic disturbance that caused the L4 cable system outage. *Geophysical Research Letters*, 26(5), 577–580. <https://doi.org/10.1029/1999GL900035>
- Clilverd, M. A., Rodger, C. J., Brundell, J. B., Dalzell, M., Martin, I., Mac Manus, D. H., & Obana, Y. (2018). Long-lasting geomagnetically induced currents and harmonic distortion observed in New Zealand during the 7–8 September 2017 disturbed Period. *Space Weather*, 16(6). <https://doi.org/10.1029/2018sw001822>
- Daldorff, L. K. S., Tóth, G., Gombosi, T. I., Lapenta, G., Amaya, J., Markidis, S., & Brackbill, J. U. (2014). Two-way coupling of a global Hall magnetohydrodynamics model with a local implicit particle-in-cell model. *Journal of Computational Physics*, 268, 236–254. <https://doi.org/10.1016/j.jcp.2014.03.009>
- De Zeeuw, D. L., Gombosi, T. I., Groth, C. P. T., Powell, K. G., & Stout, Q. F. (2000). An adaptive MHD method for global space weather simulations. *IEEE Transactions on Plasma Science*, 28(6), 1956–1965. <https://doi.org/10.1109/27.902224>
- De Zeeuw, D. L., Sazykin, S., Wolf, R. A., Gombosi, T. I., Ridley, A. J., & Tóth, G. (2004). Coupling of a global MHD code and an inner magnetospheric model: Initial results. *Journal of Geophysical Research*, 109(A12), A12219. <https://doi.org/10.1029/2003JA010366>
- Dimmock, A. P., Rosenqvist, L., Hall, J. O., Viljanen, A., Yordanova, E., Honkonen, I., et al. (2019). The GIC and geomagnetic response over Fennoscandia to the 7–8 September 2017 geomagnetic storm. *Space Weather*, 17, 2018SW002132. <https://doi.org/10.1029/2018SW002132>
- Dimmock, A. P., Rosenqvist, L., Welling, D. T., Viljanen, A., Honkonen, I., Boynton, R. J., & Yordanova, E. (2020). On the regional variability of  $dB/dt$  and its significance to GIC. *Space Weather*, 18(8), e2020SW002497. <https://doi.org/10.1029/2020SW002497>
- Engelbreton, M. J., Steinmetz, E. S., Posch, J. L., Pilipenko, V. A., Moldwin, M. B., Connors, M. G., et al. (2019). Night time magnetic perturbation events observed in Arctic Canada: 2. Multiple-instrument observations. *Journal of Geophysical Research: Space Physics*, 124(9), 7459–7476. <https://doi.org/10.1029/2019JA026797>
- Forsyth, C., Rae, I. J., Coxon, J. C., Freeman, M. P., Jackman, C. M., Gjerloev, J., & Fazakerley, A. N. (2015). A new technique for determining substorm onsets and phases from indices of the electrojet (sophie). *Journal of Geophysical Research: Space Physics*, 120(12), 10592–10606. <https://doi.org/10.1002/2015JA021343>
- Freeman, M. P., Forsyth, C., & Rae, I. J. (2019). The Influence of substorms on extreme rates of change of the surface horizontal magnetic field in the United Kingdom. *Space Weather*, 17, 827–844. <https://doi.org/10.1029/2018SW002148>
- Haiducek, J. D., Welling, D. T., Morley, S. K., Ganushkina, N. Y., & Chu, X. (2020). Using multiple signatures to improve accuracy of substorm identification. *Journal of Geophysical Research: Space Physics*, 125(4), e2019JA027559. <https://doi.org/10.1029/2019JA027559>
- Ingham, M., & Rodger, C. J. (2018). Telluric field variations as drivers of variations in cathodic protection potential on a natural gas pipeline in New Zealand. *Space Weather*, 16(9), 1396–1409. <https://doi.org/10.1029/2018SW001985>
- Ivannikova, E., Kruglyakov, M., Kuvshinov, A., Rastätter, L., & Pulkkinen, A. (2018). Regional 3-d modeling of ground electromagnetic field due to realistic geomagnetic disturbances. *Space Weather*, 16, 476–500. <https://doi.org/10.1002/2017SW001793>
- Juusola, L., Vanhamäki, H., Viljanen, A., & Smirnov, M. (2020). Induced currents due to 3d ground conductivity play a major role in the interpretation of geomagnetic variations. *Annals of Geophysics*, 38(5), 983–998. <https://doi.org/10.5194/angeo-38-983-2020>
- Kappenman, J. G. (2003). Storm sudden commencement events and the associated geomagnetically induced current risks to ground-based systems at low-latitude and midlatitude locations. *Space Weather*, 1(3). <https://doi.org/10.1029/2003sw000009>
- Kelbert, A. (2020). The role of global/regional earth conductivity models in natural geomagnetic hazard mitigation. *Surveys in Geophysics*, 41(115), 115–166. <https://doi.org/10.1007/s10712-019-09579-z>
- Kelbert, A., & Lucas, G. M. (2020). Modified GIC estimation using 3-D Earth conductivity. *Space Weather*, 18(8), e2020SW002467. <https://doi.org/10.1029/2020SW002467>
- Kwagala, N. K., Hesse, M., Moretto, T., Tenfjord, P., Norgren, C., Tóth, G., et al. (2020). Validating the space weather modeling framework (SWMF) for applications in northern Europe. *Journal of Space Weather and Space Climate*, 10, 33. <https://doi.org/10.1051/swsc/2020034>
- Lehtinen, M., & Pirjola, R. (1985). Currents produced in earthed conductor networks by geomagnetically-induced electric fields. *Annales Geophysicae*, 3(4), 479.
- Love, J. J., Lucas, G. M., Bedrosian, P. A., & Kelbert, A. (2019). Extreme-value geoelectric amplitude and polarization across the northeast United States. *Space Weather*, 17(3), 379–395. <https://doi.org/10.1029/2018SW002068>
- Marshalko, E., Kruglyakov, M., Kuvshinov, A., Murphy, B. S., Rastätter, L., Ngwira, C., & Pulkkinen, A. (2020). Exploring the influence of lateral conductivity contrasts on the storm time behavior of the ground electric field in the eastern United States. *Space Weather*, 18(3), e2019SW002216. <https://doi.org/10.1029/2019SW002216>
- Mukhopadhyay, A., Welling, D. T., Liemohn, M. W., Ridley, A. J., Chakraborty, S., & Anderson, B. J. (2020). Conductance model for extreme events: Impact of auroral conductance on space weather forecasts. *Space Weather*, 18, 1–39. <https://doi.org/10.1029/2020SW002551>
- Ngwira, C. M., Pulkkinen, A., Kuznetsova, M. M., & Gloer, A. (2014). Modeling extreme "Carrington-type" space weather events using three-dimensional global MHD simulations. *Journal of Geophysical Research: Space Physics*, 119(6), 4456–4474. <https://doi.org/10.1002/2013JA019661>
- Ngwira, C. M., Pulkkinen, A., Leila Mays, M., Kuznetsova, M. M., Galvin, A. B., Simunac, K., et al. (2013). Simulation of the 23 July 2012 extreme space weather event: What if this extremely rare CME was earth directed? *Space Weather*, 11(12), 671–679. <https://doi.org/10.1002/2013SW000990>
- Ngwira, C. M., Pulkkinen, A., Bernabeu, E., Eichner, J., Viljanen, A., & Crowley, G. (2015). Characteristics of extreme geoelectric fields and their possible causes: Localized peak enhancements. *Geophysical Research Letters*, 42(17), 6916–6921. <https://doi.org/10.1002/2015GL065061>
- Ngwira, C. M., Sibeck, D., Silveira, M. V. D., Georgiou, M., Weygand, J. M., Nishimura, Y., & Hampton, D. (2018). A study of intense local  $dB/dt$  variations during two geomagnetic storms. *Space Weather*, 16(6), 676–693. <https://doi.org/10.1029/2018sw001911>
- Piersanti, M., Di Matteo, S., Carter, B. A., Currie, J., & D'Angelo, G. (2019). Geoelectric field evaluation during the September 2017 geomagnetic storm: MA.I.GIC. model. *Space Weather*, 17(8), 1241–1256. <https://doi.org/10.1029/2019SW002202>
- Powell, K. G., Roe, P. L., Linde, T. J., Gombosi, T. I., & De Zeeuw, D. L. (1999). A solution-adaptive upwind scheme for ideal magnetohydrodynamics. *Journal of Computational Physics*, 154(2), 284–309. <https://doi.org/10.1006/jcp.1999.6299>
- Pulkkinen, A., Bernabeu, E., Eichner, J., Viljanen, A., & Ngwira, C. (2015). Regional-scale high-latitude extreme geoelectric fields pertaining to geomagnetically induced currents. *Earth Planets and Space*, 67(1), 93. <https://doi.org/10.1186/s40623-015-0255-6>
- Pulkkinen, A., Bernabeu, E., Thomson, A., Viljanen, A., Pirjola, R., Boteler, D., et al. (2017). Geomagnetically induced currents: Science, engineering, and applications readiness. *Space Weather*, 15(7), 828–856. <https://doi.org/10.1002/2016SW001501>
- Pulkkinen, A., Kuznetsova, M., Ridley, A., Raeder, J., Vapirev, A., Weimer, D., et al. (2011). Geospace environment modeling 2008–2009 challenge: Ground magnetic field perturbations. *Space Weather*, 9(2), 02004. <https://doi.org/10.1029/2010SW000600>

- Pulkkinen, A., Lindahl, S., Viljanen, A., & Pirjola, R. (2005). Geomagnetic storm of 29–31 October 2003: Geomagnetically induced currents and their relation to problems in the Swedish high-voltage power transmission system. *Space Weather*, 3(8). <https://doi.org/10.1029/2004SW000123>
- Pulkkinen, A., Pirjola, R., Boteler, D., Viljanen, A., & Yegorov, I. (2001a). Modelling of space weather effects on pipelines. *Journal of Applied Geophysics*, 48(4), 233–256. [https://doi.org/10.1016/S0926-9851\(01\)00109-4](https://doi.org/10.1016/S0926-9851(01)00109-4)
- Pulkkinen, A., Rastätter, L., Kuznetsova, M., Hesse, M., Ridley, A., Raeder, J., et al. (2010). Systematic evaluation of ground and geostationary magnetic field predictions generated by global magnetohydrodynamic models. *Journal of Geophysical Research*, 115(A3), A03206. <https://doi.org/10.1029/2009JA014537>
- Pulkkinen, A., Rastätter, L., Kuznetsova, M., Singer, H., Balch, C., Weimer, D., et al. (2013). Community-wide validation of geospace model ground magnetic field perturbation predictions to support model transition to operations. *Space Weather*, 11(6), 369–385. <https://doi.org/10.1002/swe.20056>
- Pulkkinen, A., Thomson, A., Clarke, E., & McKay, A. (2003). April 2000 geomagnetic storm: Ionospheric drivers of large geomagnetically induced currents. *Annals of Geophysics*, 21(3), 709–717. <https://doi.org/10.5194/angeo-21-709-2003>
- Pulkkinen, A., Viljanen, A., Pajunpää, K., & Pirjola, R. (2001b). Recordings and occurrence of geomagnetically induced currents in the Finnish natural gas pipeline network. *Journal of Applied Geophysics*, 48(4), 219–231. [https://doi.org/10.1016/S0926-9851\(01\)00108-2](https://doi.org/10.1016/S0926-9851(01)00108-2)
- Rae, I. J., Kabin, K., Lu, J. Y., Rankin, R., Milan, S. E., Fenrich, F. R., et al. (2010). Comparison of the open-closed separatrix in a global magnetospheric simulation with observations: The role of the ring current. *Journal of Geophysical Research*, 115(A8). <https://doi.org/10.1029/2009JA015068>
- Raeder, J., Wang, Y. L., Fuller-Rowell, T. J., & Singer, H. J. (2001). Global simulation of magnetospheric space weather effects of the bastille day storm. *Solar Physics*, 204, 323–337. <https://doi.org/10.1023/A:1014228230714>
- Ridley, A. J., Gombosi, T. I., & Dezeuw, D. L. (2004). Ionospheric control of the magnetosphere: Conductance. *Annals of Geophysics*, 22(2), 567–584. <https://doi.org/10.5194/angeo-22-567-2004>
- Rosenqvist, L., & Hall, J. O. (2019). Regional 3-D modeling and verification of geomagnetically induced currents in Sweden. *Space Weather*, 17(1), 27–36. <https://doi.org/10.1029/2018SW002084>
- Rosenqvist, L., Opgenoorth, H., Buchert, S., McCrea, I., Amm, O., & Lathuillere, C. (2005). Extreme solar-terrestrial events of October 2003: High-latitude and Cluster observations of the large geomagnetic disturbances on 30 October. *Journal of Geophysical Research*, 110(A9). <https://doi.org/10.1029/2004JA010927>
- Sazykin, S., Wolf, R. A., Spiro, R. W., Gombosi, T. I., De Zeeuw, D. L., & Thomsen, M. F. (2002, may). Interchange instability in the inner magnetosphere associated with geosynchronous particle flux decreases. *Geophysical Research Letters*, 29(10), 88–1. <https://doi.org/10.1029/2001GL014416>
- Shao, X., Guzdar, P. N., Milikh, G. M., Papadopoulos, K., Goodrich, C. C., Sharma, A., et al. (2002). Comparing ground magnetic field perturbations from global MHD simulations with magnetometer data for the 10 January 1997 magnetic storm event. *Journal of Geophysical Research*, 107(A8), SMP11-1–SMP11-10. <https://doi.org/10.1029/2000JA000445>
- Smith, A. W., Freeman, M. P., Rae, I. J., & Forsyth, C. (2019). The influence of sudden commencements on the rate of change of the surface horizontal magnetic field in the United Kingdom. *Space Weather*, 17(11), 1605–1617. <https://doi.org/10.1029/2019SW002281>
- Sorathia, K. A., Merkin, V. G., Panov, E. V., Zhang, B., Lyon, J. G., Garretson, J., et al. (2020). Ballooning-interchange instability in the near-Earth plasma sheet and auroral beads: Global magnetospheric modeling at the limit of the MHD approximation. *Geophysical Research Letters*, 47(14), e2020GL088227. <https://doi.org/10.1029/2020GL088227>
- Tanskanen, E. I. (2009). A comprehensive high-throughput analysis of substorms observed by IMAGE magnetometer network: Years 1993–2003 examined. *Journal of Geophysical Research*, 114(5), A05204. <https://doi.org/10.1029/2008JA013682>
- Toffoletto, F., Sazykin, S., Spiro, R., & Wolf, R. (2003). Inner magnetospheric modeling with the rice convection model. *Space Science Reviews*, 107, 175–196. [https://doi.org/10.1023/A:102553200804710.1007/978-94-007-1069-6\\_19](https://doi.org/10.1023/A:102553200804710.1007/978-94-007-1069-6_19)
- Tóth, G., Meng, X., Gombosi, T. I., & Rastätter, L. (2014). Predicting the time derivative of local magnetic perturbations. *Journal of Geophysical Research: Space Physics*, 119(1), 310–321. <https://doi.org/10.1002/2013JA019456>
- Tóth, G., Sokolov, I. V., Gombosi, T. I., Chesney, D. R., Clauer, C. R., De Zeeuw, D. L., et al. (2005, 12). Space weather modeling framework: A new tool for the space science community. *Journal of Geophysical Research*, 110(A12), 12226. <https://doi.org/10.1029/2005JA011126>
- Tóth, G., van der Holst, B., Sokolov, I. V., De Zeeuw, D. L., Gombosi, T. I., Fang, F., et al. (2012). Adaptive numerical algorithms in space weather modeling. *Journal of Computational Physics*, 231(3), 870–903. <https://doi.org/10.1016/j.jcp.2011.02.006>
- Tsurutani, B. T., Hajra, R., Echer, E., & Gjerloev, J. W. (2015). Extremely intense (SML ≤ −2500 nT) substorms: Isolated events that are externally triggered?. *Annals of Geophysics*, 33(5), 519–524. <https://doi.org/10.5194/angeo-33-519-2015>
- Viljanen, A. (1997). The relation between geomagnetic variations and their time derivatives and implications for estimation of induction risks. *Geophysical Research Letters*, 24(6), 631–634. <https://doi.org/10.1029/97GL00538>
- Viljanen, A., Nevanlinna, H., Pajunpää, K., & Pulkkinen, A. (2001). Time derivative of the horizontal geomagnetic field as an activity indicator. *Annals of Geophysics*, 19(9), 1107–1118. <https://doi.org/10.5194/angeo-19-1107-2001>
- Viljanen, A., & Pirjola, R. (2017). Influence of spatial variations of the geoelectric field on geomagnetically induced currents. *Journal of Space Weather Space and Climate*, 7, A22. <https://doi.org/10.1051/swsc/2017024>
- Viljanen, A., Pulkkinen, A., Pirjola, R., Pajunpää, K., Posio, P., & Koistinen, A. (2006a). Recordings of geomagnetically induced currents and a nowcasting service of the Finnish natural gas pipeline system. *Space Weather*, 4(10). <https://doi.org/10.1029/2006SW000234>
- Viljanen, A., Tanskanen, E. I., & Pulkkinen, A. (2006b). Relation between substorm characteristics and rapid temporal variations of the ground magnetic field. *Annals of Geophysics*, 24(2), 725–733. <https://doi.org/10.5194/angeo-24-725-2006>
- Welling, D. (2019). Magnetohydrodynamic models of B and their use in GIC estimates. In J. L. Gannon, A. Swidinsky, & Z. Xu (Eds.), *Geomagnetically induced currents from the sun to the power grid (in press)* (pp. 43–65). American Geophysical Union (AGU). <https://doi.org/10.1002/9781119434412.ch3>
- Welling, D. T., Anderson, B. J., Crowley, G., Pulkkinen, A. A., & Rastätter, L. (2017). Exploring predictive performance: A reanalysis of the geospace model transition challenge. *Space Weather*, 15(1), 192–203. <https://doi.org/10.1002/2016SW001505>
- Werner, A. L. E., Yordanova, E., Dimmock, A. P., & Temmer, M. (2019, 2). Modeling the multiple CME interaction event on 6–9 September 2017 with WSA-ENLIL+Conc. *Space Weather*, 17(2), 357–369. <https://doi.org/10.1029/2018SW001993>
- Wik, M., Pirjola, R., Lundstedt, H., Viljanen, A., Wintoft, P., & Pulkkinen, A. (2009). Space weather events in July 1982 and October 2003 and the effects of geomagnetically induced currents on Swedish technical systems. *Annals of Geophysics*, 27(4), 1775–1787. <https://doi.org/10.5194/angeo-27-1775-2009>
- Yu, Y., & Ridley, A. J. (2008). Validation of the space weather modeling framework using ground-based magnetometers. *Space Weather*, 6(5). <https://doi.org/10.1029/2007SW000345>